

A microscopic image showing numerous rod-shaped bacteria, likely E. coli, expressing a green fluorescent protein (GFP). The bacteria are distributed across the field of view, with some appearing as single cells and others in small clusters. The background is a uniform green color.

# QUANTITATIVE SYSTEMS BIOLOGY FOR ENGINEERING ORGANISMS AND PATHWAYS

EDITED BY : Hilal Taymaz-Nikerel and Alvaro R. Lara

PUBLISHED IN: Frontiers in Bioengineering and Biotechnology



# frontiers

## Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-829-0

DOI 10.3389/978-2-88919-829-0

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)



# QUANTITATIVE SYSTEMS BIOLOGY FOR ENGINEERING ORGANISMS AND PATHWAYS

Topic Editors:

**Hilal Taymaz-Nikerel**, Bogazici University, Turkey

**Alvaro R. Lara**, Universidad Autónoma Metropolitana-Cuajimalpa, Mexico

Studying organisms as a whole for potential metabolic(ally) engineering of organisms for production of (bio)chemicals is essential for industrial biotechnology. To this end, integrative analysis of different –omics measurements (transcriptomics, proteomics, metabolomics, fluxomics) provides invaluable information. Combination of experimental top-down and bottom-up approaches with powerful analytical tools/techniques and mathematical modeling, namely (quantitative) systems biology, currently making the state of art of this discipline, is the only practice that would improve our understanding for the purpose.

The use of high-throughput technologies induced the required development of many bioinformatics tools and mathematical methods for the integration of obtained data. Such research is significant since compiling information from different levels of a living system and connecting them is not an easy task. In particular, construction of dynamic models for product improvement has been one of the goals of many research groups.

In this Research Topic, we summarize and bring a general review of the most recent and relevant contributions in quantitative systems biology applied in metabolic modeling perspective. We want to make special emphasis on the techniques that can be widely implemented in regular scientific laboratories and in those works that include theoretical presentations.

With this Research Topic we discuss the importance of applying systems biology approaches for finding metabolic engineering targets for the efficient production of the desired biochemical integrating information from genomes and networks to industrial production. Examples and perspectives in the design of new industrially relevant chemicals, e.g. increased titer/productivity/yield of (bio)chemicals, are welcome. Addition to the founded examples, potential new techniques that would frontier the research will be part of this topic. The significance of multi ‘omics’ approaches to understand/uncover the pathogenesis/mechanisms of metabolic diseases is also one of the main topics.

**Citation:** Taymaz-Nikerel, H., Lara, A. R., eds. (2016). Quantitative Systems Biology for Engineering Organisms and Pathways. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-829-0

# Table of Contents

- 04 Editorial: Quantitative Systems Biology for Engineering Organisms and Pathways**  
Hilal Taymaz-Nikerel and Alvaro R. Lara
- 06 Lean-proteome strains – next step in metabolic engineering**  
Kaspar Valgepea, Karl Peebo, Kaarel Adamberg and Raivo Vilu
- 10 CO<sub>2</sub> – intrinsic product, essential substrate, and regulatory trigger of microbial and mammalian production processes**  
Bastian Blombach and Ralf Takors
- 21 Production of cinnamic and p-hydroxycinnamic acids in engineered microbes**  
Alejandra Vargas-Tah and Guillermo Gosset
- 31 ColE1-plasmid production in Escherichia coli: mathematical simulation and experimental validation**  
Inga Freudenau, Petra Lutter, Ruth Baier, Martin Schleeef, Hanna Bednarz, Alvaro R. Lara and Karsten Niehaus
- 45 Shikimic acid production in Escherichia coli: from classical metabolic engineering strategies to omics applied to improve its production**  
Juan Andrés Martínez, Francisco Bolívar and Adelfo Escalante
- 61 Fluorescent reporter libraries as useful tools for optimizing microbial cell factories: a review of the current methods and applications**  
Frank Delvigne, Hélène Pêcheux and Cédric Tarayre
- 69 Modifying Yeast Tolerance to Inhibitory Conditions of Ethanol Production Processes**  
Luis Caspeta, Tania Castillo and Jens Nielsen
- 84 Systems Biology Approaches to Understand Natural Products Biosynthesis**  
Cuauhtemoc Licona-Cassani, Pablo Cruz-Morales, Angel Manteca, Francisco Barona-Gomez, Lars K. Nielsen and Esteban Marcellin
- 92 Systems Biology of Microbial Exopolysaccharides Production**  
Ozlem Ates
- 108 Genome-Wide Transcriptional Response of Saccharomyces cerevisiae to Stress-Induced Perturbations**  
Hilal Taymaz-Nikerel, Ayca Cankorur-Cetinkaya and Betul Kirdar





# Editorial: Quantitative Systems Biology for Engineering Organisms and Pathways

Hilal Taymaz-Nikerel<sup>1\*</sup> and Alvaro R. Lara<sup>2</sup>

<sup>1</sup> Department of Chemical Engineering, Bogazici University, Istanbul, Turkey, <sup>2</sup> Departamento de Procesos y Tecnología, Universidad Autónoma Metropolitana-Cuajimalpa, Mexico City, Mexico

**Keywords:** metabolic engineering, systems biology, industrial biotechnology, -omics, regulation

## The Editorial on the Research Topic

### Quantitative Systems Biology for Engineering Organisms and Pathways

The biological production of chemicals has gained interest due to its contribution to greener and sustainable processes. Discovering the metabolic capacities of microorganisms shows ascending promise. The understanding of the metabolism and its complex interactions within the process environment is crucial to successfully apply and design cell factories. With the advances of high throughput measurements of -omic levels in a cell, it is possible to decipher the knowledge at different biological processes as a whole. Combining the gained experimental data with computational methods, namely, employing systems biology, allows the development of new bio-products and new cell factories.

This Research Topic of *Frontiers in Bioengineering and Biotechnology* includes six reviews, two mini-reviews, an opinion, and an original research article. Mainly, the importance of applying systems biology approaches for finding metabolic engineering targets is discussed. Valgepea et al. explained their opinion on the potential of proteome optimization in contribution to feasible development of bioprocesses. Delvigne et al. summarized the importance of using fluorescent reporter libraries for the optimization of microbial production under bioreactor environments. They focused on the current status of this technique in terms of methods and applications. Martínez et al. reported the rational design examples of *Escherichia coli* strains for the production of shikimic acid, a precursor aromatic compound in the synthesis of a drug, which is efficient against diverse viruses, including H5N1 and H1N1. In addition, they discussed the challenging tasks required for further improving the overproducing strains using global transcriptomic analyses. Vargas-Tah and Gosset explored the microbial production of two other aromatic intermediates: cinnamic and *p*-hydroxycinnamic acids. The approaches in metabolic engineering of various microorganisms to optimize the usage of raw material and increase the efficiencies of these products were explained. Licona-Cassani et al. provided recent efforts on the systems biology studies in actinomycetes, which are important pathogens and valuable sources of antibiotics, in relation to optimize the production of bioactive natural products. Ates evaluated another example on the application of omics technologies in her review, the production of microbial exopolysaccharides. In addition to these studies on the application of systems biology tools for industrial-scale production, Freudenau et al. presented a dynamic mathematical model constructed to understand the factors affecting the production of plasmid DNA as a pharmaceutical gene vector.

Blombach and Takors reviewed the impacts of carbon dioxide/bicarbonate levels on the physiological, production, metabolism, and regulation processes in microbial and mammalian cultures, a relevant issue still poorly understood. Caspeta et al. focused on the lignocellulosic production of ethanol in *Saccharomyces cerevisiae* and reviewed the information available on the inhibitory conditions during such processes and related stress response mechanisms of the

## OPEN ACCESS

### Edited and reviewed by:

Pierre De Meyts,  
De Meyts R&D Consulting, Belgium

### \*Correspondence:

Hilal Taymaz-Nikerel  
hilal.taymaz@gmail.com

### Specialty section:

This article was submitted to  
Systems Biology,  
a section of the journal  
*Frontiers in Bioengineering and  
Biotechnology*

**Received:** 09 February 2016

**Accepted:** 22 February 2016

**Published:** 08 March 2016

### Citation:

Taymaz-Nikerel H and Lara AR (2016)  
Editorial: Quantitative Systems  
Biology for Engineering Organisms  
and Pathways.  
*Front. Bioeng. Biotechnol.* 4:22.  
doi: 10.3389/fbioe.2016.00022

cells. Understanding the reprogramming of cellular functions in response to stress environments is crucial, and Taymaz-Nikerel et al. reviewed the current knowledge gained mainly by transcriptomic studies carried out in *S. cerevisiae*. Furthermore, they addressed the requirements of construction of a quantitative whole cell model, which is the ultimate goal of systems biology.

We expect that the developments in this field will continue to increase, eventually yielding quantitative predictive models, which will be useful in many areas.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Taymaz-Nikerel and Lara. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use,

## AUTHOR CONTRIBUTIONS

Both authors participated equally in the preparation of this contribution, have read, and approved the final manuscript.

## ACKNOWLEDGMENTS

This work is supported by The Scientific and Technological Research Council of Turkey (TUBITAK) through project no. 114C062.

*distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*



# Lean-proteome strains – next step in metabolic engineering

Kaspar Valgepea<sup>1</sup>, Karl Peebo<sup>1,2</sup>, Kaarel Adamberg<sup>1,3</sup> and Raivo Vilu<sup>1,2\*</sup>

<sup>1</sup> Competence Center of Food and Fermentation Technologies, Tallinn, Estonia

<sup>2</sup> Department of Chemistry, Tallinn University of Technology, Tallinn, Estonia

<sup>3</sup> Department of Food Processing, Tallinn University of Technology, Tallinn, Estonia

\*Correspondence: raivo@kbfi.ee

## Edited by:

Hilal Taymaz Nikerel, Bogazici University, Turkey

## Reviewed by:

Steve Van Dien, Genomatica, USA

**Keywords:** strain engineering, recombinant cells, absolute quantitative proteomics, genome engineering, *Escherichia coli*, genome reduction, chassis cells, whole-cell model

Rapid development of high-throughput -omics (e.g., proteomics) and genetic engineering technologies together with an array of new metabolic modeling tools during this century has led to the emergence of new fields of biological research termed systems biology and synthetic biology. The successful exploitation of these developments is evidenced by the creation of increasing number of genetically engineered recombinant cells with superior characteristics (Jantama et al., 2008; Becker et al., 2011) or totally novel functions (Nakamura and Whited, 2003; Yim et al., 2011; Paddon et al., 2013) for diverse sectors such as chemicals and healthcare (Huang et al., 2012; Lee et al., 2012; Sun and Alper, 2014). However, there exists a significant gap in bioprocess performance between studies of the literature and the requirements for an industrially feasible bioprocess for chemical production (Van Dien, 2013). Overall bioprocess performance [productivity (gram/liter/hour), titer (gram/liter) etc.] has to be increased further for successful industrial-scale commercialization to drive the shift from fossil fuel to bioprocess-based chemical production and cost-effective production of novel drugs (Van Dien, 2013). Hence, there is great need for novel approaches addressing these key challenges in chemical and healthcare sectors.

## POTENTIAL OF PROTEOME OPTIMIZATION

With this opinion, we propose that a novel approach of proteome optimization carries a substantial potential for addressing the aforementioned challenges

in bioprocess development. That potential arises from the fact that cells express proteins not essential (e.g., flagellar, heat or acid stress proteins) for growth under well-controlled optimal conditions, typically realized in biotechnological processes. This leads to non-efficient use of protein synthesis capacity (translation machinery) and energy for bioprocesses. As translation capacity is believed to be one of the growth-limiting factors, at least in the bacterium *Escherichia coli* (Klump et al., 2013), synthesis of non-essential proteins sequesters ribosomes potentially lowering the synthesis capacity of target molecule production. Thus removing the expression burden of non-essential proteins, i.e., creation of lean-proteome strains, could enable to specifically manipulate the allocation of ribosomes for higher synthesis of proteins leading to increased target molecule production. Optimization of the cellular proteome through experimental testing of strains with optimized expression of non-essential proteins and inclusion of protein synthesis capacity constraints in metabolic modeling could open a new avenue for the creation of superior cell factories.

Initial experimental confirmation of the potential of optimization of the layer of protein synthesis capacity for increasing the maximum specific growth rate ( $\mu_{\max}$ ) of cells comes from two studies of *E. coli* investigating the effects of heterologous protein expression on  $\mu$  (Scott et al., 2010; Bienick et al., 2014). Both studies show for several heterologous proteins (e.g., LacZ, eGFP) that increasing their expression has a linear negative effect on  $\mu$ . Their data suggest that for expression of every 1% of

heterologous protein per dry cell weight,  $\mu$  decreases by ~3%. It would be sensible to assume that a similar correlation would exist for the opposite case – decreasing the fraction of non-essential proteins by 1% would lead to an increase in  $\mu$  by ~3%. Our proposal is also supported by two studies of *Bacillus subtilis* showing that reducing the expression load of proteins non-essential under bioprocess conditions by ~9% fraction from the total proteome through the deletion of the flagellar/motility regulator gene *sigD* leads to a ~30% increase of both  $\mu_{\max}$  and biomass yield (Fischer and Sauer, 2005; Muntel et al., 2014). Further support comes from recent experiments of D'Souza et al. (2014), which show that deletion of single amino acid, vitamin, or nucleobase biosynthesis genes from *E. coli* results in higher  $\mu_{\max}$  compared to the wild-type strain when both strains are grown on medium containing the amino acid, vitamin, or nucleobase that the deletion strain was auxotrophic for. These observations are consistent with earlier chemostat studies with *B. subtilis* (Zamenhof and Eichhorn, 1967) and *E. coli* (Dykhuizen, 1978) where mutants impaired in tryptophan biosynthesis demonstrate significant fitness advantages in the presence of tryptophan relative to prototrophic cells. More importantly, D'Souza et al. (2014) show that deleting genes with higher protein expression cost leads to a greater growth advantage.

The results presented above suggest that proteome resource optimization through decreasing the fraction of non-essential proteins could lead to faster growth and thus also to better bioprocess performance.



For instance, target molecule productivity could be increased in growth-coupled production processes by enabling faster growth at the same expression level(s) of target molecule production-related proteins. On the other hand, recombinant protein titers could be significantly elevated by allocating more proteome resources for target protein expression at the expense of lower synthesis of non-essential proteins even at the same  $\mu$  and/or protein synthesis rate.

## REDUCED-GENOME APPROACHES

A conceptually similar approach of creating reduced-genome strains for industrial purposes has been applied in few cases before (Pósfai et al., 2006; Mizoguchi et al., 2008; Unthan et al., 2014; Xue et al., 2014). However, these efforts concentrated on reducing the genome and neglected the effects of gene deletions on the cellular proteome. The approach of deleting large chunks of the genome, instead of specific genes, based on gene function and not on protein abundance was probably responsible for the observed minor positive effects on cellular growth and target molecule production. While the latter studies focused on large-scale genome reduction, experimental technologies enabling more targeted and accurate engineering of strains with reduced load of gene expression have recently emerged. Hence, now the successful execution of the concept of targeted optimization of the layer of protein synthesis capacity is feasible due to the recent rapid progress in proteome-wide absolute quantitative proteomics (Arike et al., 2012; Ahrné et al., 2013; Wiśniewski et al., 2014) and high-throughput genome engineering technologies [e.g., Multiplexed Automated Genome Engineering (MAGE; Wang et al., 2009), trackable multiplex recombineering (TRMR; Warner et al., 2010)]. Thus, the time is ripe to design and create lean-proteome strains possibly leading to superior bioprocess performance.

## CHALLENGES WITH PROTEOME OPTIMIZATION

The main challenge with creating lean-proteome strains is hitting the correct genes/proteins, i.e., genes, which deletion does not lead to detrimental effects. This is a serious concern even in the most studied bacterium *E. coli* since functions for a third of its proteins are still

unknown (Keseler et al., 2013) while only ~300 proteins are considered essential for *E. coli* ([http://ecoliwiki.net/colipedia/index.php/Essential\\_genes](http://ecoliwiki.net/colipedia/index.php/Essential_genes)). It is important to point out that knowing functions/essentiality for more proteins is not the objective *per se* – it is actually more important to know the functions/essentiality of the proteins with the biggest translational burden (abundance  $\times$  length), as their deletion presumably leads to stronger effects. The good news here is that for many organisms, the proteome mass (a good proxy for length) distribution follows the Pareto principle – ~20% of proteins make up ~80% of the proteome mass (Ghaemmaghami et al., 2003; Maier et al., 2011; Schmidt et al., 2011; Valgepea et al., 2013). Thus, instead of targeting hundreds of genes/genome areas like in the reduced-genome approach described above, one could theoretically greatly increase the key metrics of bioprocess performance (titer, yield, productivity; Van Dien, 2013) by deleting as few as ~10 non-essential genes with the highest translational burden in *E. coli* (in total 7% of proteome; Valgepea et al., 2013) and substituting the “freed” 7% of the total proteome with target molecule-related proteins. Importantly, current mass-spectrometric techniques of absolute proteome quantification (Arike et al., 2012; Ahrné et al., 2013; Wiśniewski et al., 2014) are accurate enough to determine the proteins with the biggest translational burden on the whole-proteome level.

## STRATEGIES OF PROTEOME OPTIMIZATION FOR CREATING LEAN-PROTEOME STRAINS

The first and most important step toward creating lean-proteome strains is absolute quantitative proteome analysis of the initial recombinant strain. Accurate characterization of the full proteome is needed for the compilation of lists of non-essential target proteins with the biggest translational burden. We propose two strategies for creating superior lean-proteome strains by targeting proteins with the biggest translational burden, currently specifically for *E. coli*:

1. The first strategy targets proteins with known functions and presumably unnecessary under optimal bioprocess

conditions, e.g., pH, temperature, oxygen tension control; defined substrate feed; stirring. These could be proteins involved in stress responses (acid, heat, and osmotic shock), alternative substrate transport and catabolism and cellular movement (flagellar).

2. The second strategy targets proteins with unknown functions with the biggest translational burden. Beneficial for both approaches is the growth screen of all the Keio collection single (Baba et al., 2006) and double deletion strains (personal communication with Prof. Hirotada Mori) that can be used to determine the genes/proteins, which should and should not be targeted.

Another important step is the experimental construction of lean-proteome strains and selection for better production strains. Instead of reducing the proteome one protein at a time, one should target tens of genes with an approach similar to MAGE (Wang et al., 2009), which constantly generates genetic heterogeneity in the pool of mutants allowing the generation of thousands of lean-proteome strains within a few days. The challenge of selecting for better production strains could be tackled by combining several screening methods. First, one could screen for fast growth as reduction of non-essential protein expression should lead to faster growth. Second, high-producing strains could be isolated using fluorescence activated cell sorting (FACS) using a sensor system based on a fluorescent readout corresponding to target molecule levels.

## POTENTIAL OF METABOLIC MODELING

Lastly, one would greatly benefit from an *in silico* metabolic model, which would enable quantitative prediction of the effects of removing non-essential proteins on target molecule production. This should be a model, which incorporates the cellular proteome with the two central features of regulation of  $\mu$  – cell geometry and cell cycle – and ties the latter to the fluxes of flux balance analysis (FBA)-type models for *in silico* analysis and design of lean-proteome strains. Recently, we have seen serious progress into this direction by the development of a novel single-cell model (Abner et al., 2013), next-generation FBA-type of genome-scale

models of metabolism and gene expression (O'Brien et al., 2013; Liu et al., 2014), and a whole-cell model (Karr et al., 2012). Surely, these models will be advanced further and hopefully they will also be able to determine which genes/proteins to delete for creating superior lean-proteome strains.

## CONCLUSION

Based on the recent rapid advances in high-throughput mutant generation and proteomics technologies together with the emerging novel whole-cell modeling approaches, we conclude that the time is ripe for the metabolic engineering community to directly focus on proteome optimization leading to the creation of lean-proteome strains with superior target molecule production characteristics.

## ACKNOWLEDGMENTS

The financial support for this work was provided by the European Regional Development Fund project EU29994 and institutional research (IUT 1927) and personal (G9192) funding of the Estonian Ministry of Education and Research.

## REFERENCES

- Abner, K., Aaviksaar, T., Adamberg, K., and Vilu, R. (2013). Single-cell model of prokaryotic cell cycle. *J. Theor. Biol.* 341C, 78–87. doi:10.1016/j.jtbi.2013.09.035
- Ahrné, E., Molzahn, L., Glatter, T., and Schmidt, A. (2013). Critical assessment of proteome-wide label-free absolute abundance estimation strategies. *Proteomics* 17, 2567–2578. doi:10.1002/pmic.201300135
- Arike, L., Valgepea, K., Peil, L., Nahku, R., Adamberg, K., and Vilu, R. (2012). Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli*. *J. Proteomics* 75, 5437–5448. doi:10.1016/j.jprot.2012.06.020
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., et al. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2, 2006.0008. doi:10.1038/msb4100050
- Becker, J., Zelder, O., Häfner, S., Schröder, H., and Wittmann, C. (2011). From zero to hero – design-based systems metabolic engineering of *Corynebacterium glutamicum* for L-lysine production. *Metab. Eng.* 13, 159–168. doi:10.1016/j.ymben.2011.01.003
- Bienick, M. S., Young, K. W., Klesmith, J. R., Detwiler, E. E., Tomek, K. J., and Whitehead, T. A. (2014). The interrelationship between promoter strength, gene expression, and growth rate. *PLoS ONE* 9:e109105. doi:10.1371/journal.pone.0109105
- D'Souza, G., Waschina, S., Pande, S., Bohl, K., Kaleta, C., and Kost, C. (2014). Less is more: selective advantages can explain the prevalent loss of biosynthetic genes in bacteria. *Evolution* 68, 2559–2570. doi:10.1111/evo.12468
- Dykhuizen, D. E. (1978). Selection for tryptophan auxotrophs of *Escherichia coli* in glucose-limited chemostats as a test of energy-conservation hypothesis. *Evolution* 32, 125–150. doi:10.2307/2407415
- Fischer, E., and Sauer, U. (2005). Large-scale in vivo flux analysis shows rigidity and suboptimal performance of *Bacillus subtilis* metabolism. *Nat. Genet.* 37, 636–640. doi:10.1038/ng1555
- Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., et al. (2003). Global analysis of protein expression in yeast. *Nature* 425, 737–741. doi:10.1038/nature02046
- Huang, C.-J., Lin, H., and Yang, X. (2012). Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J. Ind. Microbiol. Biotechnol.* 39, 383–399. doi:10.1007/s10295-011-1082-9
- Jantama, K., Zhang, X., Moore, J. C., Shanmugam, K. T., Svoronos, S. A., and Ingram, L. O. (2008). Eliminating side products and increasing succinate yields in engineered strains of *Escherichia coli*. *Biotechnol. Bioeng.* 101, 881–893. doi:10.1002/bit.22005
- Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival, B., et al. (2012). A whole-cell computational model predicts phenotype from genotype. *Cell* 150, 389–401. doi:10.1016/j.cell.2012.05.044
- Keseler, I. M., Mackie, A., Peralta-Gil, M., Santos-Zavaleta, A., Gama-Castro, S., Bonavides-Martínez, C., et al. (2013). EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res.* 41, D605–D612. doi:10.1093/nar/gks1027
- Klumpp, S., Scott, M., Pedersen, S., and Hwa, T. (2013). Molecular crowding limits translation and cell growth. *Proc. Natl. Acad. Sci. U.S.A.* 110, 16754–16759. doi:10.1073/pnas.1310377110
- Lee, J. W., Na, D., Park, J. M., Lee, J., Choi, S., and Lee, S. Y. (2012). Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat. Chem. Biol.* 8, 536–546. doi:10.1038/nchembio.970
- Liu, J. K., O'Brien, E. J., Lerman, J. A., Zengler, K., Palsson, B. Ø., and Feist, A. M. (2014). Reconstruction and modeling protein translocation and compartmentalization in *Escherichia coli* at the genome-scale. *BMC Syst. Biol.* 8:110. doi:10.1186/s12918-014-0110-6
- Maier, T., Schmidt, A., Güell, M., Kühner, S., Gavin, A.-C., Aebersold, R., et al. (2011). Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Mol. Syst. Biol.* 7, 511. doi:10.1038/msb.2011.38
- Mizoguchi, H., Sawano, Y., Kato, J., and Mori, H. (2008). Superpositioning of deletions promotes growth of *Escherichia coli* with a reduced genome. *DNA Res.* 15, 277–284. doi:10.1093/dnares/dsn019
- Muntel, J., Fromion, V., Goelzer, A., Maass, S., Mader, U., Buttner, K., et al. (2014). Comprehensive absolute quantification of the cytosolic proteome of *Bacillus subtilis* by data independent, parallel fragmentation in liquid chromatography/mass spectrometry (LC/MSE). *Mol. Cell. Proteomics* 13, 1008–1019. doi:10.1074/mcp.M113.032631
- Nakamura, C. E., and Whited, G. M. (2003). Metabolic engineering for the microbial production of 1,3-propanediol. *Curr. Opin. Biotechnol.* 14, 454–459. doi:10.1016/j.copbio.2003.08.005
- O'Brien, E. J., Lerman, J. A., Chang, R. L., Hyduke, D. R., and Palsson, B. Ø. (2013). Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol. Syst. Biol.* 9, 693. doi:10.1038/msb.2013.52
- Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., et al. (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496, 528–532. doi:10.1038/nature12051
- Pósfai, G., Plunkett, G., Fehér, T., Frisch, D., Keil, G. M., Umenhoffer, K., et al. (2006). Emergent properties of reduced-genome *Escherichia coli*. *Science* 312, 1044–1046. doi:10.1126/science.1126439
- Schmidt, A., Beck, M., Malmström, J., Lam, H., Claassen, M., Campbell, D., et al. (2011). Absolute quantification of microbial proteomes at different states by directed mass spectrometry. *Mol. Syst. Biol.* 7, 510. doi:10.1038/msb.2011.37
- Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression: origins and consequences. *Science* 330, 1099–1102. doi:10.1126/science.1192588
- Sun, J., and Alper, H. S. (2014). Metabolic engineering of strains: from industrial-scale to lab-scale chemical production. *J. Ind. Microbiol. Biotechnol.* doi:10.1007/s10295-014-1539-8
- Unthan, S., Baumgart, M., Radek, A., Herbst, M., Siebert, D., Brühl, N., et al. (2014). Chassis organism from *Corynebacterium glutamicum* – a top-down approach to identify and delete irrelevant gene clusters. *Biotechnol. J.* doi:10.1002/biot.201400041
- Valgepea, K., Adamberg, K., Seiman, A., and Vilu, R. (2013). *Escherichia coli* achieves faster growth by increasing catalytic and translation rates of proteins. *Mol. Biosyst.* 2344–2358. doi:10.1039/c3mb70119k
- Van Dien, S. J. (2013). From the first drop to the first truckload: commercialization of microbial processes for renewable chemicals. *Curr. Opin. Biotechnol.* 24, 1061–1068. doi:10.1016/j.copbio.2013.03.002
- Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., Forest, C. R., et al. (2009). Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 460, 894–898. doi:10.1038/nature08187
- Warner, J. R., Reeder, P. J., Karimpour-Fard, A., Woodruff, L. B. A., and Gill, R. T. (2010). Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. *Nat. Biotechnol.* 28, 856–862. doi:10.1038/nbt.1653
- Wiśniewski, J. R., Hein, M. Y., Cox, J., and Mann, M. (2014). A “proteomic ruler” for protein copy number and concentration estimation without spike-in standards. *Mol. Cell. Proteomics* 13, 3497–3506. doi:10.1074/mcp.M113.037309
- Xue, X., Wang, T., Jiang, P., Shao, Y., Zhou, M., Zhong, L., et al. (2014). The MEGA (Multiple Essential Genes Assembling) deletion and replacement method for genome reduction in *Escherichia coli*. *ACS Synth. Biol.* doi:10.1021/sb500324p
- Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A., Boldt, J., et al. (2011). Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat. Chem. Biol.* 7, 445–452. doi:10.1038/nchembio.580

Zamenhof, S., and Eichhorn, H. H. (1967). Study of microbial evolution through loss of biosynthetic functions – establishment of defective mutants. *Nature* 26, 456–458. doi:10.1038/216456a0

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 January 2015; accepted: 22 January 2015; published online: 06 February 2015.

Citation: Valgepea K, Peebo K, Adamberg K and Vilu R (2015) Lean-proteome strains – next step in metabolic engineering. *Front. Bioeng. Biotechnol.* 3:11. doi: 10.3389/fbioe.2015.00011

This article was submitted to Systems Biology, a section of the journal *Frontiers in Bioengineering and Biotechnology*.

Copyright © 2015 Valgepea, Peebo, Adamberg and Vilu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# CO<sub>2</sub> – intrinsic product, essential substrate, and regulatory trigger of microbial and mammalian production processes

**Bastian Blombach and Ralf Takors\***

*Institute of Biochemical Engineering, University of Stuttgart, Stuttgart, Germany*

## OPEN ACCESS

### Edited by:

Hilal Taymaz Nikerel,  
Bogazici University, Turkey

### Reviewed by:

Anshu Bhardwaj,  
Council of Scientific and Industrial  
Research, India  
Peter Neubauer,  
Technische Universität  
Berlin, Germany  
Antonino Baez,  
Autonomous University of  
Puebla, Mexico

### \*Correspondence:

Ralf Takors,  
Institute of Biochemical Engineering,  
University of Stuttgart,  
Allmandring 31, 70569  
Stuttgart, Germany  
takors@ibvt.uni-stuttgart.de

### Specialty section:

This article was submitted to  
Systems Biology, a section of the  
journal *Frontiers in Bioengineering  
and Biotechnology*

**Received:** 22 May 2015

**Accepted:** 13 July 2015

**Published:** 03 August 2015

### Citation:

Blombach B and Takors R (2015)  
CO<sub>2</sub> – intrinsic product, essential  
substrate, and regulatory trigger of  
microbial and mammalian  
production processes.  
*Front. Bioeng. Biotechnol.* 3:108.  
doi: 10.3389/fbioe.2015.00108

Carbon dioxide formation mirrors the final carbon oxidation steps of aerobic metabolism in microbial and mammalian cells. As a consequence, CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> dissociation equilibria arise in fermenters by the growing culture. Anaplerotic reactions make use of the abundant CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels for refueling citric acid cycle demands and for enabling oxaloacetate-derived products. At the same time, CO<sub>2</sub> is released manifold in metabolic reactions via decarboxylation activity. The levels of extracellular CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> depend on cellular activities and physical constraints such as hydrostatic pressures, aeration, and the efficiency of mixing in large-scale bioreactors. Besides, local CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels might also act as metabolic inhibitors or transcriptional effectors triggering regulatory events inside the cells. This review gives an overview about fundamental physicochemical properties of CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> in microbial and mammalian cultures effecting cellular physiology, production processes, metabolic activity, and transcriptional regulation.

**Keywords:** bicarbonate, carbon dioxide, production process, regulation, carboxylation, decarboxylation

## Introduction

One of the most decisive decisions which needs to be made when developing novel bioprocesses is whether the final process will run under anaerobic or aerobic conditions. While severely reduced investment costs speak in favor of anaerobic production, expected productivities and intracellular energy availabilities are drivers for aerobic approaches. Anaerobic metabolism yields at two net ATP produced in glycolysis per glucose while aerobic counterparts may achieve >12 ATPs. This net ATP yield even represents a conservative estimation considering true ATP per oxygen (P/O) ratios of 1:1.3 which are lower than theoretical maxima of 2–3. Consequently, aerobic processes are often the first choice if ATP-challenging product formation with maximum cell-specific formation rates is targeted.

Carbon dioxide (CO<sub>2</sub>) is the inevitable product of respiration processes and as such always present in aerobic bioprocesses. This holds also true for the production of commodities, fine chemicals, or therapeutic proteins using microbes or mammalian cells. While therapeutic proteins and fine chemicals are typically produced in bioreactor of 5–20 m<sup>3</sup> scale, the production of commodities is usually performed in 50–500 m<sup>3</sup> size – or even larger. As an intrinsic property, partial CO<sub>2</sub> pressures of these scales differ significantly from those found in lab-scale. This phenomenon is the inherent consequence of high absolute pressures and poor mixing conditions in large-scale bioreactors (Takors, 2012).

CO<sub>2</sub> and its hydrated counterpart HCO<sub>3</sub><sup>-</sup> may not only serve as substrate or product for carboxylating and decarboxylating reactions, the species may also alter physicochemical properties of proteins, acidify the internal pH, and regulate virulence and toxin production in pathogens (Follonier et al., 2013). Therefore, CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> can interact with cellular metabolism and can even create complex transcriptional responses. CO<sub>2</sub> not only freely diffuses through the cellular membrane (Gutknecht et al., 1977), it may also accumulate in the same (Jones and Greenfield, 1982; Kuriyama et al., 1993; Bothun et al., 2004), thus increasing its permeability and fluidity which finally leads in the potentially lethal “anesthesia effect” (Isenschmid et al., 1995).

The fact that high pCO<sub>2</sub> levels are used to sterilize food (Ballestra et al., 1996; Spilimbergo and Bertucco, 2003; Garcia-Gonzalez et al., 2007) anticipates that elevated pCO<sub>2</sub> are not likely to improve the performance of microbial or mammalian production processes. Instead, as it will be shown, high pCO<sub>2</sub> levels often coincide with the deterioration of the bioprocess performance. Consequently, thorough scale-up studies should preferably consider the analysis of pCO<sub>2</sub> impacts to ensure an equally good performance in large-scale compared to lab-scale expectations. This is especially true for the establishment of novel bioprocesses which are the result of systems metabolic engineering studies performed in lab-scale.

This contribution aims at reviewing fundamental properties, sources, and impacts of CO<sub>2</sub> for microbial and mammalian production processes. It yields at bringing together the major puzzle pieces of how CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> interacts with producer cells. It will show that a lot has already been done – but still not everything is fully understood. This holds especially true for regulation of cellular metabolism where CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> apparently serves as an underestimated trigger so far.

## Fundamentals – Physicochemical Properties and Mass Transfer

Carbon dioxide (CO<sub>2</sub>, molar weight: 44.01 g/mol) is a colorless, odorless gas of linear molecular shape with a melting point at –56.6°C. It is present in the Earth atmosphere as a trace compound currently showing levels of about 400 ppm with the tendency of steadily increase (<http://co2now.org/>).

The water solubility can be described applying Henry's law.

$$H_{\text{CO}_2} = \frac{c_{\text{CO}_2,\text{L}}}{p_{\text{CO}_2}} \left[ \frac{\text{mmol}}{\text{L bar}} \right] \quad (1)$$

with  $c_{\text{CO}_2,\text{L}}$  and  $p_{\text{CO}_2}$  coding for the equilibrium values of the molar concentration of dissolved CO<sub>2</sub> in the liquid L and the related partial CO<sub>2</sub> pressure, respectively. For pure water at 25°C the Henry coefficient  $H_{\text{CO}_2} = 34.5 \text{ mmol/barL}$  is given (Stumm and Morgan, 1995). Using the Van't Hoff correlation

$$\frac{d \ln H_{\text{CO}_2}}{dT} = \frac{\Delta H^0}{RT^2} \Rightarrow H_{\text{CO}_2}(T) = \ln K - \frac{\Delta H^0}{RT} \quad (2)$$

the temperature dependency of the equilibrium constant (here: Henry-coefficient  $H_{\text{CO}_2}$ ) can be estimated with the standard

enthalpy change of the reaction  $\Delta H^0$ , the universal gas constant  $R$ , and the absolute temperature  $T$  as shown. Noteworthy,  $K$  codes for an integration constant that can be derived from reference data e.g., at 25°C. Using equation (2),  $H_{\text{CO}_2}(T = 20^\circ\text{C}) = 40 \text{ mmol/barL}$  and  $H_{\text{CO}_2}(T = 37^\circ\text{C}) = 25 \text{ mmol/barL}$  can be calculated. Decreasing Henry coefficients [as defined by (1)] mirror reducing gas solubility with rising temperature – a typical phenomenon for dissolved gasses at the given temperature range.

Besides temperature, CO<sub>2</sub> solubility is also affected by electrolyte concentrations. Following the empirical Sechenov (1889) approach individual contributions of ion strength can be considered to estimate the resulting solubility of a gas in the salt-containing liquid (Noorman et al., 1992). However, the composition of fermentation media is often complex and changes steadily during the course of cultivation. Product and by-product formation, substrate consumption, and the addition of titrating agents are the reasons. Therefore, the most pragmatic approach is to measure CO<sub>2</sub> solubility in real cultivation media. Own experimental observations show that real  $H_{\text{CO}_2}$  values [according to (1)] are often increased, may be even doubled, compared to values for pure water (unpublished data).

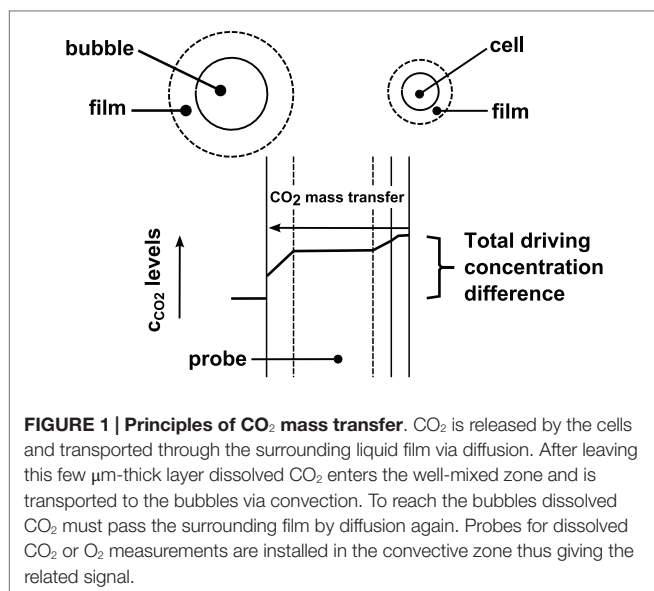
Applying typical operating conditions, microbial or mammalian cultivations release exhaust gas with volumetric CO<sub>2</sub> fractions of 5–25%. For a conservative estimation, one can assume equilibrium conditions between gas and liquid with  $H_{\text{CO}_2}$  values for pure water at 37°C. Then dissolved CO<sub>2</sub> levels  $c_{\text{CO}_2,\text{L}}$  are likely to range between 75 and 375 mg/L. For instance, Blombach et al. (2013) measured pCO<sub>2</sub> levels of about 160 mbar (about 360 mg/L) at the end of an aerated (0.1 vvm) 1.5 bar pressured, stirred batch cultivation with 5 g<sub>CDW</sub> *Corynebacterium glutamicum* per L. Increasing the aeration to 3 vvm reduced the pCO<sub>2</sub> to 40 mbar (about 90 mg/L). Similar values were observed by Buchholz et al. (2014b). By contrast, maximum dissolved oxygen concentrations under atmospheric conditions will typically result at 7.5–8 mg/L (again depending on medium composition). Consequently, dissolved CO<sub>2</sub> levels outcompete dissolved O<sub>2</sub> levels by far. This finding may be even more pronounced if mass transport characteristics are considered (Figure 1).

Figure 1 shows that maximum  $c_{\text{CO}_2,\text{L}}$  levels are found in the proximate microenvironment of the cells. By trend, probes for dissolved gas measurement observe lower levels. This is different compared to dissolved oxygen where cells face the lowest levels along the mass transfer path.

While dissolved carbon dioxide levels may achieve high inhibiting values during the fermentation course, starting conditions might be limiting instead. Assuming equilibrium between inlet aeration and the liquid 0.5 mg<sub>CO<sub>2</sub></sub>/L is present. Noteworthy this low value is likely to persist if too high aeration (with low concentrated CO<sub>2</sub>) strips out new metabolically produced CO<sub>2</sub>. Consequently anaplerotic reactions may be limited by substrate (HCO<sub>3</sub><sup>-</sup> / CO<sub>2</sub>) supply (see Section “Metabolic Release and Incorporation”) finally resulting at reduced cell growth.

By analogy to oxygen transfer, the CO<sub>2</sub> transfer rate  $CTR$  (mmol/Lh) can be described according to the following:

$$CTR = k_L a_{\text{CO}_2} (c_{\text{CO}_2,\text{L}}^* - c_{\text{CO}_2,\text{L}}) \quad (3)$$



with  $k_L a_{\text{CO}_2}$  coding for the CO<sub>2</sub> mass transfer coefficient (1/h),  $c_{\text{CO}_2, \text{L}}$  for the dissolved CO<sub>2</sub> concentration at equilibrium following Henry's law (mmol/L) and  $c_{\text{CO}_2, \text{L}}$  representing the measured concentration (mmol/L).

Measuring true  $k_L a_{\text{CO}_2}$  values in praxis is somewhat challenging. One approach is to assume  $CTR = CER$ , i.e., carbon dioxide emission rate  $CER$  equals the CO<sub>2</sub> stripping rate  $CTR$ . By balancing flows of aeration and exhaust gas, related values should be accessible and  $k_L a_{\text{CO}_2}$  can be derived accordingly. Nevertheless, this approach reveals its drawback when mammalian cell cultures are balanced. Here, the exhaust gas signal is a superposition of biological activity and CO<sub>2</sub> addition for titration. Alternatively,  $k_L a_{\text{CO}_2}$  could be estimated from  $k_L a_{\text{O}_2}$  according to the following:

$$k_L a_{\text{CO}_2} = k_L a_{\text{O}_2} \sqrt{\frac{D_{\text{CO}_2}}{D_{\text{O}_2}}} \quad (4)$$

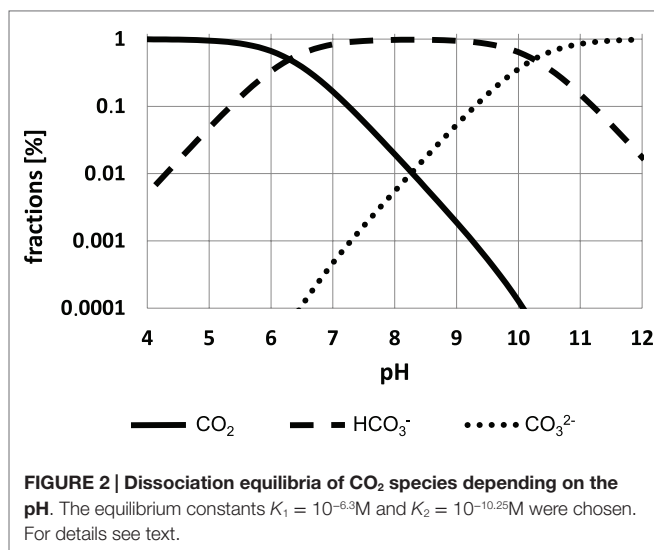
Equation (4) results from Higbie's penetration theory Higbie (1935) and Danckwerts surface renewable model Danckwerts (1951). Apparently, the mass transfer coefficient for CO<sub>2</sub> is proportionally linked to the ratio of the diffusion coefficients  $D$  for CO<sub>2</sub> and O<sub>2</sub> in water. As  $k_L a_{\text{O}_2}$  values are relatively easy to measure, the approach offers a straightforward access to  $k_L a_{\text{CO}_2}$ . However, CO<sub>2</sub> transfer differs fundamentally from O<sub>2</sub> transport because dissociation characteristics have been taken into account (see **Figure 2**).

CO<sub>2</sub> dissociates in water depending on pH as follows:



Because the equilibrium of CO<sub>2</sub> dissociation is far on the side of the anhydride (99.8%), concentrations of the carbonic acid H<sub>2</sub>CO<sub>3</sub> are low not exceeding one digit micromolar ranges at typical cultivation conditions. Consequently, the apparent equilibrium constant  $K_1$  (Bailey and Ollis, 1986):

$$K_1 = \frac{[\text{H}^+][\text{HCO}_3^-]}{[\text{CO}_2] + [\text{H}_2\text{CO}_3]} \approx \frac{[\text{H}^+][\text{HCO}_3^-]}{[\text{CO}_2]} = 10^{-6.3} \text{ M} \quad (6)$$



is formulated and completed by Bailey and Ollis (1986) as follows:

$$K_2 = \frac{[\text{H}^+][\text{CO}_3^{2-}]}{[\text{HCO}_3^-]} = 10^{-10.25} \text{ M} \quad (7)$$

One may safely assume that (de-) protonating reactions of formula (5) are very fast. However, formation and dissociation of carbonic acid from CO<sub>2</sub> are suspected to limit the total equilibration process.  $k_1$  and  $k_{-1}$  were estimated as 0.03 1/s and 20 1/s, respectively (Bailey and Ollis, 1986).

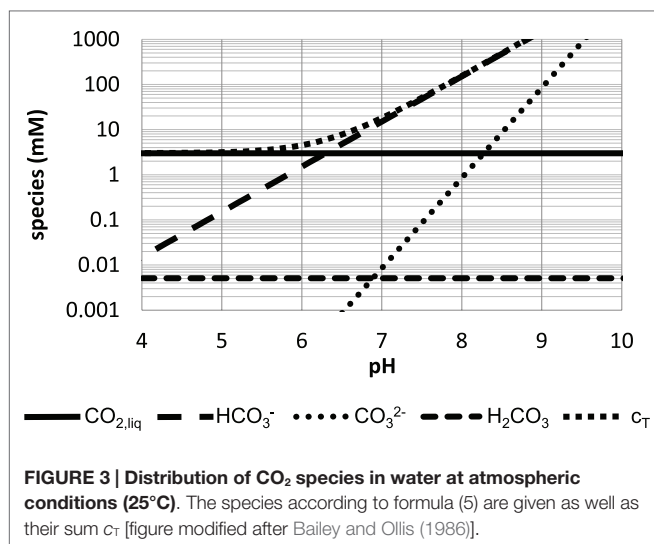
At typical cultivation conditions (pH 7), 83.3% of the CO<sub>2</sub> species are present as HCO<sub>3</sub><sup>-</sup>, only 16.7% as CO<sub>2</sub>. Hence HCO<sub>3</sub><sup>-</sup> is about five-fold higher concentrated than CO<sub>2</sub>. This statement not only holds for the cultivation medium, but it should also be valid for intracellular conditions because cells aim at maintaining their intracellular pH at about this level.

**Figure 3** underpins that the full consideration of the individual species CO<sub>2</sub>, HCO<sub>3</sub><sup>-</sup>, and CO<sub>3</sub><sup>2-</sup> is crucial to get accurate values for total CO<sub>2</sub>  $C_T$  dissolved in the fermentation suspension. Recently, Buchholz et al. (2014a) outlined that ignoring the anions leads to a carbon gap of about 20% during the first hours of fermentation. Noteworthy, the dissolved CO<sub>2</sub> level is not dependent on pH (see **Figure 3**). According to Henry's law only partial pressure (and salt conditions) may effect  $c_{\text{CO}_2}$ . Hence, large-scale bioreactors which have high hydrostatic pressures of 1–1.5 bar possess higher dissolved CO<sub>2</sub> levels than comparable laboratory systems. This not only induces regulatory responses in the cells but also affects the buffering capacity of the large-scale suspension. Due to increased CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels, pH buffering is severely increased in large scale compared to lab fermentations.

## Metabolic Release and Incorporation

Metabolism of all living organisms is equipped with a set of carboxylases incorporating CO<sub>2</sub> or bicarbonate (HCO<sub>3</sub><sup>-</sup>) into organic molecules and decarboxylases releasing CO<sub>2</sub> in the environment. Consequently, these fundamental reactions are directly involved in and/or interconnect anabolism,





catabolism, and energy metabolism of the cell. Especially, the phosphoenolpyruvate–pyruvate–oxaloacetate node comprises an organism-specific configuration of carboxylating (e.g., pyruvate carboxylase (PCx), PEP carboxylase, acetyl-CoA carboxylase) and decarboxylating (e.g., PEP carboxykinase; malic enzyme, oxaloacetate decarboxylase; pyruvate dehydrogenase complex, pyruvate:quinone oxidoreductase) reactions (Figure 4) which are of major importance for the carbon flux distribution in the central metabolism. For instance during sugar catabolism anaplerotic C3 (phosphoenolpyruvate (PEP)/pyruvate), carboxylation and decarboxylation of pyruvate to acetyl-CoA are essentially required to maintain TCA flux whereas gluconeogenesis relies on C4 (oxaloacetate/malate) decarboxylation (Sauer and Eikmanns, 2005). Another example is pyruvate decarboxylase of yeast which is the key enzyme in ethanol fermentation and is essentially required to maintain a balanced metabolism in mineral media containing glucose as sole carbon source (Pronk et al., 1996).

Carboxylases catalyzing the thermodynamically expensive assimilation of CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> have been classified regarding their physiological function into autotrophic, assimilatory, biosynthetic, anaplerotic, and redox balancing enzymes (Erb, 2011). Currently, six pathways for CO<sub>2</sub> fixation have been identified: the reductive pentose phosphate (Calvin–Benson) cycle, the reductive acetyl-CoA (Wood–Ljungdahl) pathway, the reductive citric acid cycle, the 3-hydroxypropionate bicycle, the dicarboxylate/4-hydroxybutyrate cycle, and the 3-hydroxypropionate/4-hydroxybutyrate cycle (Erb, 2011; Fuchs, 2011). Additionally, to increase the carbon fixation rate, novel synthetic pathways have been proposed (Bar-Even et al., 2010). Plants, algae and phototrophic prokaryotes possess ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) quantitatively the most abundant enzyme in the biosphere and the key enzyme in autotrophic CO<sub>2</sub> fixation by the Calvin–Benson cycle (Miziorko and Lorimer, 1983; Hartman and Harpel, 1994; Erb, 2011). Notably, the Wood–Ljungdahl pathway is the only one which fixes CO<sub>2</sub> and simultaneously generates ATP by conversion of acetyl-CoA to acetate (Fuchs, 2011) rendering this route attractive and promising for CO<sub>2</sub>-based microbial production purposes (Dürre and Eikmanns, 2015).

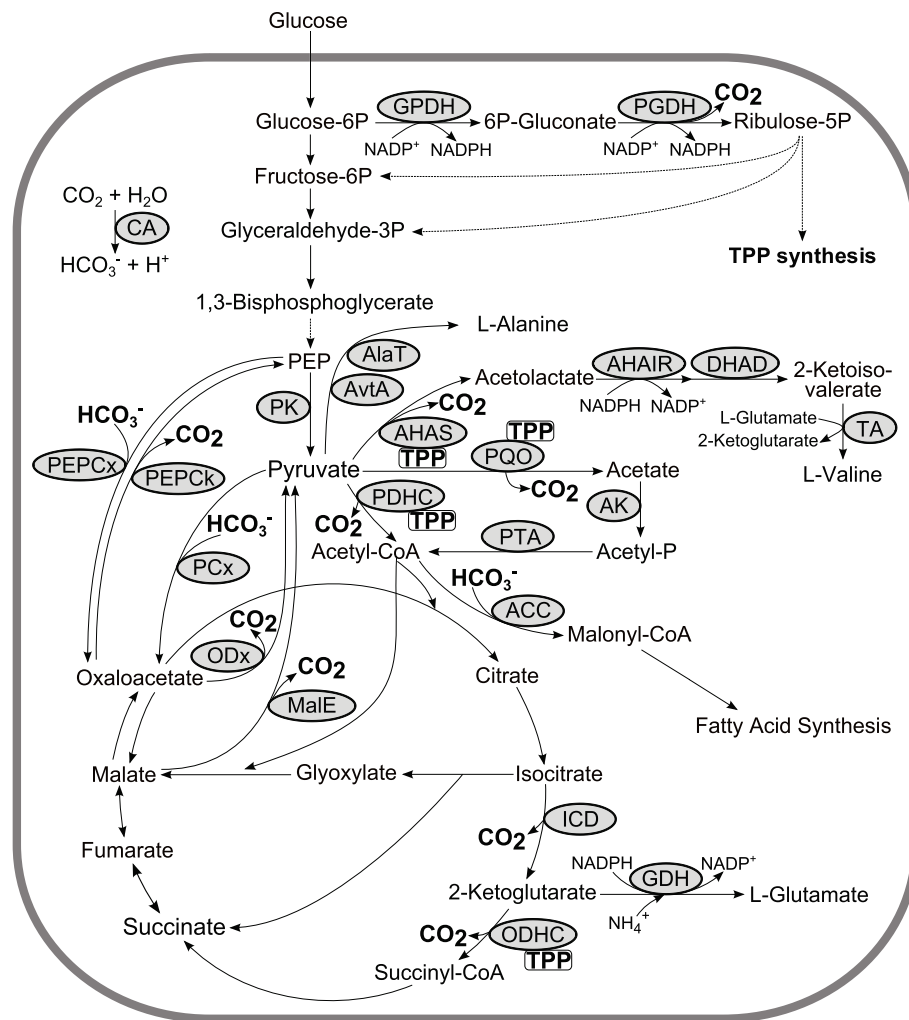
Owing to the metabolic key functions carboxylases and decarboxylases fulfill, relevant microbial, mammalian, and plant enzymes have been biochemically characterized and their regulation analyzed [e.g., Miziorko and Lorimer (1983), Chollet et al. (1996), Hanson and Reshef (1997), Nikolau et al. (2003), Sauer and Eikmanns (2005), and Jitrapakdee et al. (2006)]. CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup>-related allosteric regulation and biochemical properties of relevant enzymes have been analyzed to some extent (Jones and Greenfield, 1982), however, our understanding is limited by far, yet. The industrially important Gram-positive *C. glutamicum* possesses the biotin-dependent PCx and the PEP carboxylase, both with Michaelis–Menten constants ( $K_M$ ) for HCO<sub>3</sub><sup>-</sup> of about 3 mM (Hanke et al., 2005; Chen et al., 2013) which is about 30-fold higher compared to the  $K_M$  (0.1 mM i.e., 4.4 mg/L) of PEP carboxylase from *Escherichia coli* as single anaplerotic enzyme (Kai et al., 1999). These differences already point to organism-specific aeration needs to install proper metabolic activity by maintaining sufficient CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> availability at fermentation start when biomass concentrations are still low (Repaske et al., 1974; Talley and Baugh, 1975) or for products requiring a high anaplerotic flux (e.g., succinate, L-lysine, and derived products).

Due to the diffusive potential of CO<sub>2</sub> and the rather slow chemical conversion of CO<sub>2</sub> to HCO<sub>3</sub><sup>-</sup> (Kern, 1960), nature has independently evolved three classes (designated as  $\alpha$ ,  $\beta$ , and  $\gamma$ ) of zinc-dependant carbonic anhydrases (CAs) which catalyze with very high turnover numbers (up to 10<sup>6</sup> s<sup>-1</sup>) the reversible hydration of CO<sub>2</sub> (Figure 4; Tashian, 1989; Tripp et al., 2001). CAs are widespread over all kingdoms of life and play a vital role in various cellular functions such as photosynthesis, ion transport, and pH homeostasis (Smith and Ferry, 2000; Merlin et al., 2003; Mitsuhashi et al., 2004). Essentially, CAs maintain adequate HCO<sub>3</sub><sup>-</sup> levels for aerobic growth under ordinary atmospheric conditions, since inactivation of CAs in several organisms such as *C. glutamicum*, *E. coli*, *Ralstonia eutropha*, *Candida albicans*, *Saccharomyces cerevisiae*, and *Aspergillus nidulans* is lethal unless the CO<sub>2</sub> content in the atmosphere is significantly increased (about 5–10%; Mitsuhashi et al., 2004; Merlin et al., 2003; Kusian et al., 2002; Götz et al., 1999; Cottier et al., 2012).

In mammals, mitochondrial respiration generates CO<sub>2</sub> as waste product which has to be actively transported from tissue to the depolluting lungs by the blood. Since HCO<sub>3</sub><sup>-</sup> is in contrast to CO<sub>2</sub> not permeable, mammalian cells are equipped with about 13 genes encoding different types of bicarbonate transporters allowing the intercellular exchange of the former species (Casey, 2006). By contrast, HCO<sub>3</sub><sup>-</sup> transport in prokaryotes has been rarely observed so far with the well-studied exception of the bicarbonate transport system of the cyanobacteria *Synechococcus* sp. strain PCC7942. Two different transport mechanisms for HCO<sub>3</sub><sup>-</sup> maintain with the combined action of CA elevated levels of CO<sub>2</sub> in the carboxysomes required for efficient carbon fixation by RubisCO (Ritchie et al., 1996; Badger and Price, 2003).

## CO<sub>2</sub> Induced growth Phenotypes

CO<sub>2</sub> is the final respiratory product and consequently inevitable in aerobic microbial and mammalian bioprocesses. In exhaust gas



**FIGURE 4 | Exemplary, a schematic overview of the central metabolism of *C. glutamicum* including carboxylases and TPP-dependent decarboxylases.** Abbreviations: ACC, acetyl-CoA carboxylase; AHAS, acetohydroxyacid synthase; AHAS, acetohydroxyacid isomeroreductase; AK, acetate kinase; AlaT, alanine aminotransferase; AvtA, valine-pyruvate aminotransferase; CA, carbonic anhydrase; DHAD, dihydroxyacid dehydratase; GDH, glutamate dehydrogenase; GPDH, glucose-6P dehydrogenase; ICD,

isocitrate dehydrogenase; MalE, malic enzyme; ODHC, 2-ketoglutarate dehydrogenase complex; ODx, oxaloacetate decarboxylase; PCx, pyruvate carboxylase; PDHC, pyruvate dehydrogenase complex; PEP, phosphoenolpyruvate; PEPCK, PEP carboxykinase; PEPCx, PEP carboxylase; PGDH, 6P-gluconate dehydrogenase; PK, pyruvate kinase; PQO, pyruvate:quinone oxidoreductase; PTA, phosphotransacetylase; TA, transaminase B; TPP, thiamine pyrophosphate.

flows the CO<sub>2</sub> fraction may rise to 15–20% depending on aeration and cellular activity. Considering that head overpressures of microbial fermentations are commonly 0.5–2 bar and 10–15 m bioreactor filling heights create hydrostatic pressures of 1–1.5 bar, p<sub>CO2</sub> could achieve maximum values of 0.1–0.6 bar at the bottom of the bioreactor. Notably, these maximum values may be reduced if aeration with fresh air is properly installed there. In principle, the scenario is similar for mammalian cultures although lowered due to reduced cell activities and smaller bioreactor sizes compared to microbial applications (p<sub>CO2</sub> at mammalian production: about 0.180 bar<sub>CO2</sub>; Zhu et al., 2005). Noteworthy, cells circulating in large-scale bioreactors experience frequently changing p<sub>CO2</sub> levels, a fact that is usually not simulated by pseudo-stationary scale-down tests.

Multiple studies have been performed for elucidating the impact of p<sub>CO2</sub> levels on microbial (Dixon and Kell, 1989) and mammalian performance [e.g., Gray et al. (1996)]. Effects on growth, biomass per substrate yields, product formation, cell division, and morphology were analyzed. These were either attributed to elevated CO<sub>2</sub> partial pressures alone or in conjunction with co-effects such as changing osmolality in the media. Observed phenotypes are individual. Nevertheless, some characteristic examples are given in the following highlighting basic kinetics of industrially interesting strains:

## Bacteria

First indications that bacteria do react on elevated dissolved CO<sub>2</sub> levels were published by Jones and Greenfield (1982). Among

others, they observed that growth of *Bacillus subtilis* was inhibited by 40% under  $p_{\text{CO}_2} = 0.17$  atm (0.172 bar). Batch studies with *E. coli* using CO<sub>2</sub>-enriched aeration revealed that the maximum growth rate was severely reduced and biomass per glucose yields increased for aeration fractions >20% of carbon dioxide (Castan et al., 2002). Baez et al. (2009) studied GFP producing *E. coli* at constant  $p_{\text{CO}_2}$  in the range of 20–300 mbar. Their results supported previous findings by measuring more than 30% reduction of the maximal growth rate  $\mu_{\text{max}}$  and doubled acetate formation under  $p_{\text{CO}_2} = 300$  mbar compared to the reference. For *C. glutamicum*, Knoll et al. (2005) investigated the growth rates  $\mu$  in overpressurized bioreactors (10 bar head pressure) during growth on glucose. They observed  $\mu > 0.3$  1/h under  $p_{\text{CO}_2} = 0.43$  bar. This finding was supported by subsequent studies with an L-lysine producing *C. glutamicum* strain (Knoll et al., 2007). Additionally, turbidostatic continuous cultivations were performed installing different  $p_{\text{CO}_2}$  levels. The growth rate of 0.58 1/h turned out to be almost constant until 0.18 bar  $p_{\text{CO}_2}$  and steadily decreased to 0.36 1/h under 0.8 bar  $p_{\text{CO}_2}$  (Bäumchen et al., 2007). In 2013, Blombach et al. studied the growth performance of *C. glutamicum* in batch cultures. While no significant growth phenotype was found installing  $p_{\text{CO}_2}$  of about 0.3 bar, low levels smaller than 50 mbar  $p_{\text{CO}_2}$  revealed 3-phase, bi-level growth kinetics of *C. glutamicum* (Blombach et al., 2013). Recently, Lopes et al. (2014) reviewed some microbial phenotypes as a result of elevated carbon dioxide levels in over-pressurized bioreactors.

## Yeast

Chen and Gutmanns (1976) reported about growth inhibition of yeast at high CO<sub>2</sub> partial pressures. They found “slight” growth inhibition using CO<sub>2</sub> aeration fractions of 40% and a severe growth decrease using 50% CO<sub>2</sub> enriched air. Later, Kuriyama et al. (1993) underlined these early findings by arguing that cell division of *S. cerevisiae* may be hampered under  $p_{\text{CO}_2} = 0.5$  atm (0.51 bar). Kuriyama et al. (1993) used chemostat approaches for studying the  $p_{\text{CO}_2}$  impact. They found that an elevated  $p_{\text{CO}_2}$  coincided with increased ethanol formation which itself may hamper process performance. *S. cerevisiae* is able to adapt to hyperbaric conditions (10 bar) provided that sufficient time for adaptations is given (Belo et al., 2003). CO<sub>2</sub> partial pressures of 0.48 bar had negligible effects on cell viability. This was also observed by Knoll et al. (2007). However, if partial pressures are increased further (0.6 bar) cell budding is hampered (Coelho et al., 2004). Indeed, a growth reduction of 25% was reported by Aguilera et al. (2005) when the CO<sub>2</sub> fraction of aeration was increased to 79% in aerobic cultivations. However, growth under anaerobic conditions was much less affected indicating that the respiratory metabolism is likely to be more influenced under high  $p_{\text{CO}_2}$  levels. This phenomenon was in the focus of recent studies. Richard et al. (2014) outlined that transient metabolic responses are triggered by CO<sub>2</sub> shifts e.g., characterized by intermediary increase of respiration rates and the excretion of ethanol and acetate.

## Fungi

Similar to bacteria and yeast, inhibition of growth (and product formation) was also observed for fungi such as *Penicillium chrysogenum* already under  $p_{\text{CO}_2} = 0.08$  atm (Jones and Greenfield,

1982). Ho and Smith (1986) specified this early observation by identifying reduced growth and penicillin formation rates using 12.6% CO<sub>2</sub> enriched air for cultivation. However, causes and consequences of high  $p_{\text{CO}_2}$  levels on growth and product formation may not be clearly identifiable. They may rather be a matter of indirect effects finally resulting in morphology changes (McIntyre and McNeil, 1998). Also Gibbs et al. (2000) pinpointed to the chemical interaction of high  $p_{\text{CO}_2}$  with precursors of penicillin biosynthesis finally deteriorating performance of *P. chrysogenum*. Nevertheless, under high levels of  $p_{\text{CO}_2}$  (installed after using 10–15% enriched influent gas) increased climbing and severely reduced penicillin production were observed (El-Sabbagh et al., 2006), not only for *P. chrysogenum* but also for cephalosporin C producing *Acremonium chrysogenum* (El-Sabbagh et al., 2008).

## Mammalian Cells (e.g., CHO)

Today, mammalian producers are typically derived from tissue cells giving Chinese hamster ovary (CHO) cells an outstanding importance for the production of therapeutic proteins (Pfizenmaier and Takors, 2015). It has been estimated that these cells experience  $p_{\text{CO}_2}$  levels of 41–72 mbar under physiological conditions (Altman and Dittmer, 1971). However, industrial production environments are likely to impose much higher  $p_{\text{CO}_2}$ , especially when processes are in the focus of ongoing intensification (Ozturk, 1996).  $p_{\text{CO}_2}$ -induced stress usually coincides with the increase of osmolality due to titration for pH control. Hence, the interaction of both effects is often in the foreground of related studies. Kimura and Miller (1996) analyzed recombinant tissue-type plasminogen activator (tPA) production with CHO cells. Under maximum  $p_{\text{CO}_2}$  of 333 mbar they observed 30% reduction of the growth rate which increased to 45% reduction in combination with high osmolality. Results of Gray et al. (1996) anticipated that an optimum for recombinant protein production exists at 40–100 mbar  $p_{\text{CO}_2}$ . Zhu et al. (2005) showed that industrial osmolality conditions (400–450 mOsm) together with typically high  $p_{\text{CO}_2}$  (180–213 mar) levels caused a 20% drop of CHO cell viability. Besides, Takuma et al. (2007) outlined that industrial  $p_{\text{CO}_2}$  values of 293 mbar reduced growth by 60% while cell-specific productivity of antibody IgG1 was almost unchanged. Additionally, there were indications that appropriate glucose limitation could compensate  $p_{\text{CO}_2}$  triggered growth reduction at “moderate” 190 mbar  $p_{\text{CO}_2}$ .

Among others, one reason for the deteriorating performance may be that protein glycosylation patterns reduce in the presence of elevated HCO<sub>3</sub><sup>-</sup> levels (Zanghi et al., 1999). Besides, DeZengotita et al. (2002) argued that glycolysis was inhibited in a dose-dependent manner when  $p_{\text{CO}_2}$  levels were studied between 66 and 333 mbar in hybridoma cells. Therefore,  $p_{\text{CO}_2}$  inhibition is not only a matter of CHO cells alone, but is observed for hybridoma and HEK293S cultures as well (Jardon and Garnier, 2003).

## CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup>-Induced Regulation

CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> not only serves as substrate or product for enzymes, but also impacts the internal pH, the fluidity and permeability of membranes, and physicochemical properties of proteins, and is regarded as signal for virulence and toxin production in pathogens



(Isenschmid et al., 1995; Stretton and Goodman, 1998; Follonier et al., 2013). Due to the multiple involvement of CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> in cellular metabolism, it seems evident that these species are directly or indirectly part of the regulatory machinery.

The human body underlies a complex CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> homeostasis with bicarbonate concentrations up to 140 mM in certain tissues (Arthurs and Sudhakar, 2005; Abuaita and Withey, 2009; Orłowski et al., 2013) representing a striking signal for pathogens invading the host. Although a direct association between CO<sub>2</sub> and virulence is missing, Park et al. (2011) found that 10% CO<sub>2</sub> stimulated aerobic growth of the human gastric pathogen *Helicobacter pylori*. CO<sub>2</sub> deprivation led to increased intracellular ppGpp levels which might indicate an involvement of the stringent response in CO<sub>2</sub>-dependent regulation of *H. pylori*'s metabolism (Park et al., 2011). In *Vibrio cholerae* bicarbonate activates the regulatory protein ToxT which in turn induces virulence gene expression (Abuaita and Withey, 2009). Another, bicarbonate sensing transcriptional regulator is the AraC-like protein RegA from the mouse enteric pathogen *Citrobacter rodentium* which in the presence of bicarbonate activates transcription of a number of virulence genes and inhibits expression of several housekeeping genes (Yang et al., 2009). *C. albicans* a fungal pathogen causing life-threatening infections in immunocompromised patients senses increased HCO<sub>3</sub><sup>-</sup> levels by the soluble adenylyl cyclase (sAC) Cyr1p which produces cAMP. Then, cAMP activates protein kinase A to trigger filamentous growth which is an important feature for adhesion and invasion of the pathogen (Klengel et al., 2005; Hall et al., 2010). Furthermore, the transcription factor Rca1p of *C. albicans* was shown to control expression of CA in response to the availability of CO<sub>2</sub> (Cottier et al., 2012). Both examples demonstrate the relevancy of a CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> signaling system for global regulation of *C. albicans*' metabolism. Regulation by bicarbonate-responsive soluble ACs seems to be more widespread across multiple kingdoms since CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup>-dependent adjustment of the intracellular cAMP level, initially found in male germ cells, was also identified in mycobacteria, eubacteria, fungi, and cyanobacteria (Chen et al., 2000; Zippin et al., 2001; Bahn and Mühlshlegel, 2006).

Although in large-scale fermentations gradients of dissolved gases occur and high CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> concentrations depending on the process and the production host arise (Hermann, 2003; Takors, 2012), only few studies investigated the effects of altered levels of these species on metabolism and regulation of industrial relevant microbial cells systematically. The already mentioned analysis of Baez et al. (2009) studied the effect of 300 mbar partial pressure on recombinant GFP producing *E. coli* not only metabolically but also on the transcriptional level. Expression analysis of 16 selected genes revealed only slight changes in transcription. Noteworthy, as response to elevated dissolved CO<sub>2</sub> the transcription of acid stress genes (*gadA*, *gadC*, and *adiA*) increased, indicating acidification of the internal pH by CO<sub>2</sub> (Baez et al., 2009).

Recently, Follonier et al. (2013) exposed *Pseudomonas putida* KT2440 to elevated pressure (up to 7 bar) associated with increased CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> concentrations in the bioreactor. They investigated the global transcriptional response by DNA microarrays. Physiology of *P. putida* KT2440 was hardly affected at increased pressure, however, significant changes in gene

transcription were observed: elevated CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels activated the heat-shock response and strongly affected expression of cell envelope genes pointing to an altered permeability/fluidity of the membrane (Follonier et al., 2013).

The genome-wide transcriptional response of *S. cerevisiae* to high CO<sub>2</sub> concentrations was analyzed in chemostat cultures under aerobic and anaerobic conditions. Accompanied with a more pronounced sensitivity of respiratory metabolism, high CO<sub>2</sub> levels in glucose-limited cultures led to 104 at least two-fold altered transcripts compared to 33 under anaerobic conditions. Interestingly, 50% of the affected transcripts under aerobic conditions encoded mitochondrial proteins such as PEP carboxykinase, PCx, and proteins involved in oxidative phosphorylation (Aguilera et al., 2005).

Recently, we investigated the effects of low (pCO<sub>2</sub> < 40 mbar) and high (pCO<sub>2</sub> ≥ 300 mbar) CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels on growth kinetics and the transcriptional response of *C. glutamicum* compared to standard conditions. Under high CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels growth kinetics were not affected albeit the biomass to substrate yield was increased. However, a complex transcriptional response involving 117 differentially expressed genes was observed. Among those, 60 genes were assigned to the complete DtxR/RipA regulon controlling iron homeostasis in *C. glutamicum*. The mutant *C. glutamicum* Δ*dtxR* showed significantly impaired growth under high CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> conditions (compared to the wildtype) but not under standard conditions. This finding underlines the relevancy of the master regulator for cell fitness under high CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels (Blombach et al., 2013). At low CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels *C. glutamicum* showed three distinct growth phases. In the mid-phase with slowest growth, *C. glutamicum* secreted L-alanine and L-valine into the medium and showed about two times higher activities of glucose-6-P dehydrogenase and 6-phosphogluconate dehydrogenase and a strong transcriptional response (>100 genes with altered expression) including increased transcription of almost all thiamine pyrophosphate (TPP) genes compared to standard conditions. We hypothesized that *C. glutamicum* counteracts the lack of CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> by triggering TPP biosynthesis for increasing the activities of TPP-dependent enzymes involved in CO<sub>2</sub> formation (Figure 4; Blombach et al., 2013).

Industrial scale cells are exposed to various gradients such as pH, substrates, and dissolved gases. To analyze the effects of oscillating CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels on the metabolism and transcriptional response of *C. glutamicum*, a novel three-compartment cascade bioreactor system was developed. pCO<sub>2</sub> gradients of 75–315 mbar at industry-relevant residence times of about 3.6 min did not significantly influence the growth kinetics but led to 66 differentially expressed genes compared to control conditions. Interestingly, the overall change in expression was directly linked to the pCO<sub>2</sub> gradients and the residence time of the cells in the scale-down device (Buchholz et al., 2014b).

## CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> Impacts Production Processes

Production processes on glycolytic substrates rely on the anaplerotic function of PCx and/or PEP carboxylase to replenish citric acid cycle intermediates that are incorporated for anabolic

demands and/or product formation. Especially, oxaloacetate-derived products such as L-lysine require a high anaplerotic flux. *C. glutamicum* is the workhorse in industrial L-lysine production and possesses PCx and PEP carboxylase. Several studies identified PCx and especially deregulated variants as most relevant to improve oxaloacetate supply since inactivation of PCx reduced and overexpression of the corresponding *pyc* gene significantly improved L-lysine formation in *C. glutamicum* (Peters-Wendisch et al., 2001; Ohnishi et al., 2002). Furthermore, inactivation of PEP carboxykinase led to an increase in L-lysine production with *C. glutamicum* (Riedel et al., 2001). Surprisingly, although great efforts have been made to tailor the biosynthetic pathway and to optimize precursor availability (Blombach and Seibold, 2010), the impact of altered CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels for aerobic L-lysine production has not been systematically investigated so far.

Apparently too low CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels may limit the *in vivo* activity of anaplerotic reactions. The combination of high aeration and low biomass concentration at the beginning of the fermentation is likely to cause retarded cell growth due to CO<sub>2</sub> over-stripping. By analogy, installing non-limiting CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels is especially important for zero-growth or resting cell bioprocesses. Examples are the synthesis of organic acids such as malate, fumarate and succinate which are formed anaerobically from oxaloacetate via the reductive arm of the citric acid cycle. Under such conditions only minor amounts of CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> are provided by the metabolism of the cell. However, elevated productivities can be achieved by sparging with CO<sub>2</sub> or adding carbonates to the medium to ensure sufficient HCO<sub>3</sub><sup>-</sup> for C3-carboxylation (Inui et al., 2004; Okino et al., 2005, 2008; Lu et al., 2009; Zelle et al., 2010; Zhang et al., 2010; Wieschalka et al., 2012). Inui et al. (2004) and Okino et al. (2005) showed that addition of NaHCO<sub>3</sub> to the medium significantly improved the glucose consumption rate and the succinate production rate with resting cells of *C. glutamicum* R. Radoš et al. (2014) demonstrated that sparging an anaerobic culture of non-growing *C. glutamicum* with CO<sub>2</sub> improved the succinate and acetate yield, respectively, both at the expense of lactate production. <sup>13</sup>C nuclear magnetic resonance analysis of labeling patterns in the end products verified the incorporation of bicarbonate and the formation of succinate mainly via the reductive arm of the citric acid cycle (Radoš et al., 2014). For a dual-phase (aerobic growth, anaerobic production) succinate production process with a recombinant *E. coli* strain, it was also shown that increasing the CO<sub>2</sub> content in the gas phase from 0 to 50% improved the biomass-specific production rate and the succinate yield significantly (Lu et al., 2009). In order to provide additional CO<sub>2</sub> and reduction equivalents for anaerobic succinate production from glucose, Litsanov et al. (2012) integrated the *fdh* gene encoding a formate dehydrogenase from *Mycobacterium vaccae* into the chromosome of an engineered *C. glutamicum* strain. Supplementation of formate increased the succinate yield by 20% mainly due to increased NADH availability. However, part of the formed CO<sub>2</sub> was incorporated into the product (Litsanov et al., 2012).

The shortage of oil resources and steadily rising oil prices has stimulated efforts to produce chemicals and fuels directly from CO<sub>2</sub>. Production of ethanol, isobutyraldehyde, and isobutanol from CO<sub>2</sub> and light was achieved using engineered photosynthetic bacteria such as *Rhodobacter capsulatus* and *Synechococcus*

*elongates* PCC7942 (Wahlund et al., 1996; Atsumi et al., 2009). Li et al. (2012) showed the feasibility of electrochemical supply of electrons to produce isobutanol and 3-methyl-1-butanol from CO<sub>2</sub> with engineered *R. eutropha* H16. However, the low productivity and final titer of such approaches and the reactor design is still a challenge for future industrial application. Alternatively, RubisCO was functionally expressed in heterotrophic *S. cerevisiae* to incorporate CO<sub>2</sub> as co-substrate improving ethanol production and reducing the formation of the by-product glycerol in chemostat cultures (Guadalupe-Medina et al., 2013). An innovative approach is the use of CO<sub>2</sub> and hydrogen-containing waste gases or synthesis gas as feedstock for the production of chemicals and fuels with acetogenic and carboxydophilic bacteria. Aerobic and anaerobic gas fermentation processes have been exploited for their biotechnological potential and commercial plants for ethanol production are already under construction (Dürre and Eikmanns, 2015).

Mammalian producer cells descent from rodents (like mouse or hamster) or human tissues. In case they are used in submerge culture they have undergone a (sometimes) tedious transition to yield at suspended producer cell lines. With this history in mind one may understand why product formation in producer cells such as CHO is often found to be strongly growth de-coupled (Altamirano et al., 2001). This fact is even exploited by temperature shift-down approaches (37 to ~30°C) to arrest cells in G1 phase finally increasing cell-specific protein production. By analogy, osmolality increase results at similar growth and product formation phenotypes (Ozturk and Palsson, 1991; Kumar et al., 2007). As outlined in the foregoing sections, elevated pCO<sub>2</sub> environments >100 mbar are likely to inhibit cell growth for CHO cultures. Consequently, therapeutic protein formation kinetics of the (typical) growth de-coupled type are not likely to be affected by high pCO<sub>2</sub> environments. Indeed, findings of Takuma et al. (2007) support this conclusion. In case growth-coupled product formation is observed, the impact of increased carbon dioxide partial pressures may be more pronounced. This holds also true for putative interactions of high CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels with the cellular membrane or the product proteins. However, more studies are necessary to investigate these individual effects.

## Conclusion

Summarizing the impacts of high CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels, the reduction of cellular growth is a typical phenomenon. Although the effects are very individual, sensitivities on high CO<sub>2</sub> partial pressures are less pronounced in bacteria than they are in fungi or mammalian producer cells. As a rule of thumb pCO<sub>2</sub> > 100 mbar marks the beginning of growth inhibition for the later.

On the other hand, too low CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels are likely to limit anaplerotic reactions inside the cells. Consequently, downstream precursors such as oxaloacetate could become limiting which affects not only cell growth but also biosynthesis of related metabolic products.

In general, transcriptional responses on high (or low) CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> are by far less studied than metabolic phenotypes. However, (maybe) surprising regulatory mechanisms are waiting to be discovered. An illustrative example is the case of *C.*

*glutamicum* that aims at counteracting CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> limitation by amplifying TPP biosynthesis, known as an essential co-factor for decarboxylating enzymes. High CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels apparently serve as an important stimulus for some pathogenic microbes to identify the host and to trigger related invasion programs. To what extent fragments or derivatives of such regulatory scenarios are also present in other cells also remains to be discovered.

Considering the application of microbes, yeasts, fungi, and mammalian cells in industrial bioreactors some particularities need to be taken into account. High CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> levels do not effect cells as a singular, isolated event. They rather occur in conjunction with changes of osmolality and pH that stimulate the cells manifold. The unequivocal identification of causes and consequences may be hampered intrinsically. Complex networking analysis is necessary to decipher details of CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> impacts. Examples are the link between CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> and productivity with morphology changes in fungi or the osmolality in CHO. On

the other side, equal pCO<sub>2</sub> levels may serve as a valuable scale-up criterion because they mirror the complex interaction of cellular activities, mixing, and mass transfer (Klinger et al., 2015). Furthermore, one should consider that CO<sub>2</sub> / HCO<sub>3</sub><sup>-</sup> stimuli occur dynamically under industrial operation conditions. Cells are circulating in large-scale production reactors thus experiencing frequently changing dissolved CO<sub>2</sub> levels. Consequently, comprehensive scale-up tests should mirror these conditions to ensure that promising novel producers will perform equally well in large scale – as they should.

## Acknowledgments

The authors gratefully acknowledge the funding of this work by the Deutsche Forschungsgemeinschaft (DFG), grant TA 241/5-1 and TA 241/5-2. This work was also supported by DFG within the funding programme Open Access Publishing.

## References

- Abuaita, B. H., and Withey, J. H. (2009). Bicarbonate induces *Vibrio cholerae* virulence gene expression by enhancing ToxT activity. *Infect. Immun.* 77, 4111–4120. doi:10.1128/IAI.00409-09
- Aguilera, J., Petit, T., De Winder, J. H., and Pronk, J. T. (2005). Physiological and genome-wide transcriptional responses of *Saccharomyces cerevisiae* to high carbon dioxide concentrations. *FEMS Yeast Res.* 5, 579–593. doi:10.1016/j.femsyr.2004.09.009
- Altamirano, C., Cairo, J. J., and Godia, F. (2001). Decoupling cell growth and product formation in Chinese hamster ovary cells through metabolic control. *Biotechnol. Bioeng.* 76, 351–360. doi:10.1002/bit.10096
- Altman, P. L., and Dittmer, D. S. (eds) (1971). *Biological Handbooks: Respiration and Circulation*. Bethesda: Federation of American Societies for Experimental Biology.
- Arthurs, G., and Sudhakar, M. (2005). Carbon dioxide transport. *Contin. Educ. Anaesth. Crit. Care Pain* 5, 207–210. doi:10.1093/bjaccp/mki050
- Atsumi, S., Higashide, W., and Liao, J. C. (2009). Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat. Biotechnol.* 27, 1177–1180. doi:10.1038/nbt.1586
- Badger, M. R., and Price, G. D. (2003). CO<sub>2</sub> concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *J. Exp. Bot.* 54, 609–622. doi:10.1093/jxb/erg076
- Baez, A., Flores, N., Bolívar, F., and Ramírez, O. T. (2009). Metabolic and transcriptional response of recombinant *Escherichia coli* to elevated dissolved carbon dioxide concentrations. *Biotechnol. Bioeng.* 104, 102–110. doi:10.1002/bit.22379
- Bahn, Y. S., and Mühlischlegel, F. A. (2006). CO<sub>2</sub> sensing in fungi and beyond. *Curr. Opin. Microbiol.* 9, 572–578. doi:10.1016/j.mib.2006.09.003
- Bailey, J. E., and Ollis, D. F. (1986). *Biochemical Engineering Fundamentals*. 2<sup>nd</sup> Edition, McGraw-Hill Chemical Engineering Series. New York, NY: McGraw-Hill.
- Ballestra, P., Da Silva, A. A., and Cuq, J. L. (1996). Inactivation of *Escherichia coli* by carbon dioxide under pressure. *J. Food Sci.* 61, 829–831. doi:10.1111/j.1365-2621.1996.tb12212.x
- Bar-Even, A., Noor, E., Lewis, N. E., and Milo, R. (2010). Design and analysis of synthetic carbon fixation pathways. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8889–8894. doi:10.1073/pnas.0907176107
- Bäumchen, C., Knoll, A., Husemann, B., Seletzky, J., Maier, B., Dietrich, C., et al. (2007). Effect of elevated dissolved carbon dioxide concentrations on growth of *Corynebacterium glutamicum* on D-glucose and L-lactate. *J. Biotechnol.* 128, 868–874. doi:10.1016/j.jbiotec.2007.01.001
- Belo, I., Pinheiro, R., and Mota, M. (2003). Fed-batch cultivation of *Saccharomyces cerevisiae* in a hyperbaric bioreactor. *Biotechnol. Prog.* 19, 665–671. doi:10.1021/bp0257067
- Blombach, B., Buchholz, J., Busche, T., Kalinowski, J., and Takors, R. (2013). Impact of different CO<sub>2</sub>/HCO<sub>3</sub><sup>-</sup> levels on metabolism and regulation in *Corynebacterium glutamicum*. *J. Biotechnol.* 168, 331–340. doi:10.1016/j.jbiotec.2013.10.005
- Blombach, B., and Seibold, G. M. (2010). Carbohydrate metabolism in *Corynebacterium glutamicum* and applications for the metabolic engineering of L-lysine production strains. *Appl. Microbiol. Biotechnol.* 86, 1313–1322. doi:10.1007/s00253-010-2537-z
- Bothun, G. D., Knutson, B. L., Berberich, J. A., Strobel, H. J., and Nokes, S. E. (2004). Metabolic selectivity and growth of *Clostridium thermocellum* in continuous culture under elevated hydrostatic pressure. *Appl. Microbiol. Biotechnol.* 65, 149–157. doi:10.1007/s00253-004-1554-1
- Buchholz, J., Graf, M., Blombach, B., and Takors, R. (2014a). Improving the carbon balance of fermentations by total carbon analyses. *Biochem. Eng. J.* 90, 162–169. doi:10.1016/j.bej.2014.06.007
- Buchholz, J., Graf, M., Freund, A., Busche, T., Kalinowski, J., Blombach, B., et al. (2014b). CO<sub>2</sub>/HCO<sub>3</sub><sup>-</sup> perturbations of simulated large scale gradients in a scale-down device cause fast transcriptional responses in *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 98, 8563–8572. doi:10.1007/s00253-014-6014-y
- Casey, J. R. (2006). Why bicarbonate? 1. *Biochem. Cell Biol.* 84, 930–939. doi:10.1139/O06-184
- Castan, A., Näsman, A., and Enfors, S.-O. (2002). Oxygen enriched air supply in *Escherichia coli* processes: production of biomass and recombinant human growth hormone. *Enzyme Microb. Technol.* 30, 847–854. doi:10.1016/S0141-0229(01)00490-2
- Chen, S. L., and Gutmann, F. (1976). Carbon dioxide inhibition of yeast growth in biomass production. *Biotechnol. Bioeng.* 18, 1455–1462. doi:10.1002/bit.260181012
- Chen, Y., Cann, M. J., Litvin, T. N., Iourgenko, V., Sinclair, M. L., Levin, L. R., et al. (2000). Soluble adenylyl cyclase as an evolutionarily conserved bicarbonate sensor. *Science* 289, 625–628. doi:10.1126/science.289.5479.625
- Chen, Z., Bommarreddy, R. R., Frank, D., Rappert, S., and Zeng, A.-P. (2013). Deregulation of feedback inhibition of phosphoenolpyruvate carboxylase for improved lysine production in *Corynebacterium glutamicum*. *Appl. Environ. Microbiol.* 80, 1388–1393. doi:10.1128/AEM.03535-13
- Chollet, R., Vidal, J., and O'Leary, M. H. (1996). Pyruvate carboxylase: a ubiquitous, highly regulated enzyme in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47, 273–298. doi:10.1146/annurev.arplant.47.1.273
- Coelho, M. A. Z., Belo, I., Pinheiro, R., Amaral, A. L., Mota, M., Coutinho, J. A. P., et al. (2004). Effect of hyperbaric stress on yeast morphology: study by automated image analysis. *Appl. Microbiol. Biotechnol.* 66, 318–324. doi:10.1007/s00253-004-1648-9
- Cottier, F., Raymond, M., Kurzai, O., Bolstad, M., Leewattanapasuk, W., Jiménez-López, C., et al. (2012). The bZIP transcription factor Rca1p is a central regulator of a novel CO<sub>2</sub> sensing pathway in yeast. *PLoS Pathog.* 8:e1002485. doi:10.1371/journal.ppat.1002485



- Danckwerts, P. V. (1951). Significance of liquid-film coefficients in gas absorption. *Ind. Eng. Chem.* 43, 1460–1467. doi:10.1021/ie50498a055
- DeZengotita, V. M., Schmelzer, A. E., and Miller, W. M. (2002). Characterization of hybridoma cell responses to elevated pCO<sub>2</sub> and osmolality: intracellular pH, cell size, apoptosis, and metabolism. *Biotechnol. Bioeng.* 77, 369–380. doi:10.1002/bit.10176
- Dixon, N. M., and Kell, D. B. (1989). The inhibition by CO<sub>2</sub> of the growth and metabolism of micro-organisms. *J. Appl. Bacteriol.* 67, 109–136. doi:10.1111/j.1365-2672.1989.tb03387.x
- Dürre, P., and Eikmanns, B. J. (2015). C1-carbon sources for chemical and fuel production by microbial gas fermentation. *Curr. Opin. Biotechnol.* 35, 63–72. doi:10.1016/j.copbio.2015.03.008
- El-Sabbagh, N., Harvey, L. M., and McNeil, B. (2008). Effects of dissolved carbon dioxide on growth, nutrient consumption, cephalosporin C synthesis and morphology of *Acremonium chrysogenum* in batch cultures. *Enzyme Microb. Technol.* 42, 315–324. doi:10.1016/j.enzmictec.2007.10.012
- El-Sabbagh, N., McNeil, B., and Harvey, L. M. (2006). Dissolved carbon dioxide effects on growth, nutrient consumption, penicillin synthesis and morphology in batch cultures of *Penicillium chrysogenum*. *Enzyme Microb. Technol.* 39, 185–190. doi:10.1016/j.enzmictec.2005.10.020
- Erb, T. J. (2011). Carboxylases in natural and synthetic microbial pathways. *Appl. Environ. Microbiol.* 77, 8466–8477. doi:10.1128/AEM.05702-11
- Follonier, S., Escapa, I. F., Fonseca, P. M., Henes, B., Panke, S., Zinn, M., et al. (2013). New insights on the reorganization of gene transcription in *Pseudomonas putida* KT2440 at elevated pressure. *Microb. Cell Fact.* 12, 30. doi:10.1186/1475-2859-12-30
- Fuchs, G. (2011). Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Annu. Rev. Microbiol.* 65, 631–658. doi:10.1146/annurev-micro-090110-102801
- García-González, L., Geeraerd, A. H., Spilimbergo, S., Elst, K., Van Ginneken, L., Debever, J., et al. (2007). High pressure carbon dioxide inactivation of micro-organisms in foods: the past, the present and the future. *Int. J. Food Microbiol.* 117, 1–28. doi:10.1016/j.jfoodmicro.2007.02.018
- Gibbs, P. A., Seviour, R. J., and Schmid, F. (2000). Growth of filamentous fungi in submerged culture: problems and possible solutions. *Crit. Rev. Biotechnol.* 20, 17–48. doi:10.1080/07388550091144177
- Götz, R., Gnann, A., and Zimmermann, F. K. (1999). Deletion of the carbonic anhydrase-like gene *NCE103* of the yeast *Saccharomyces cerevisiae* causes an oxygen-sensitive growth defect. *Yeast* 15, 855–864. doi:10.1002/(SICI)1097-0061(199907)15:10A<855::AID-YEA425>3.0.CO;2-C
- Gray, D. R., Chen, S., Howarth, W., Inlow, D., and Maiorella, B. L. (1996). CO<sub>2</sub> in large-scale and high-density CHO cell perfusion culture. *Cytotechnology* 22, 65–78. doi:10.1007/BF00353925
- Guadalupe-Medina, V., Wisselink, H. W., Luttik, M. A., de Hulster, E., Daran, J.-M., Pronk, J. T., et al. (2013). Carbon dioxide fixation by Calvin-Cycle enzymes improves ethanol yield in yeast. *Biotechnol. Biofuels* 6, 125. doi:10.1186/1754-6834-6-125
- Gutknecht, J., Bisson, M. A., and Tosteson, F. C. (1977). Diffusion of carbon dioxide through lipid bilayer membranes: effects of carbonic anhydrase, bicarbonate, and unstirred layers. *J. Gen. Physiol.* 69, 779–794. doi:10.1085/jgp.69.6.779
- Hall, R. A., de Sordi, L., MacCallum, D. M., Topal, H., Eaton, R., Bloor, J. W., et al. (2010). CO<sub>2</sub> acts as a signalling molecule in populations of the fungal pathogen *Candida albicans*. *PLoS Pathog.* 6:e1001193. doi:10.1371/journal.ppat.1001193
- Hanke, P. D., Sinskey, A. J., and Willis, L. B. (2005). *Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacterium*. U.S. Patent No 6,965,021.
- Hanson, R. W., and Reshef, L. (1997). Regulation of phosphoenolpyruvate carboxykinase (GTP) gene expression. *Annu. Rev. Biochem.* 66, 581–611. doi:10.1146/annurev.biochem.66.1.581
- Hartman, F. C., and Harpel, M. R. (1994). Structure, function, regulation, and assembly of D-ribulose-1,5-bisphosphate carboxylase/oxygenase. *Annu. Rev. Biochem.* 63, 197–234. doi:10.1146/annurev.bi.63.070194.001213
- Hermann, T. (2003). Industrial production of amino acids by coryneform bacteria. *J. Biotechnol.* 104, 155–172. doi:10.1016/S0168-1656(03)00149-4
- Higbie, R. (1935). The rate of absorption of a pure gas into a still liquid during short periods of exposure. *Trans. Am. Inst. Chem. Eng.* 31, 364–389.
- Ho, C. S., and Smith, M. D. (1986). Effect of dissolved carbon dioxide on penicillin fermentations: mycelial growth and penicillin production. *Biotechnol. Bioeng.* 28, 668–677. doi:10.1002/bit.260280506
- Inui, M., Murakami, S., Okino, S., Kawaguchi, H., Vertès, A. A., and Yukawa, H. (2004). Metabolic analysis of *Corynebacterium glutamicum* during lactate and succinate productions under oxygen deprivation conditions. *J. Mol. Microbiol. Biotechnol.* 7, 182–196. doi:10.1159/000079827
- Isenschmid, A., Marison, I. W., and von Stockar, U. (1995). The influence of pressure and temperature of compressed CO<sub>2</sub> on the survival of yeast cells. *J. Biotechnol.* 39, 229–237. doi:10.1016/0168-1656(95)00018-L
- Jardon, M., and Garnier, A. (2003). pH, pCO<sub>2</sub>, and temperature effect on R-adenovirus production. *Biotechnol. Prog.* 19, 202–208. doi:10.1021/bp025585a
- Jitrapakdee, S., Vidal-Puig, A., and Wallace, J. C. (2006). Anaplerotic roles of pyruvate carboxylase in mammalian tissues. *Cell. Mol. Life Sci.* 63, 843–854. doi:10.1007/s00018-005-5410-y
- Jones, R. P., and Greenfield, P. F. (1982). Effect of carbon dioxide on yeast growth and fermentation. *Enzyme Microb. Technol.* 4, 210–223. doi:10.1016/0141-0229(82)90034-5
- Kai, Y., Matsumura, H., Inoue, T., Terada, K., Nagara, Y., Yoshinaga, T., et al. (1999). Three-dimensional structure of phosphoenolpyruvate carboxylase: a proposed mechanism for allosteric inhibition. *Proc. Natl. Acad. Sci. U.S.A.* 96, 823–828. doi:10.1073/pnas.96.3.823
- Kern, D. M. (1960). The hydration of carbon dioxide. *J. Chem. Educ.* 37, 14. doi:10.1021/ed037p14
- Kimura, R., and Miller, W. M. (1996). Effects of elevated pCO<sub>2</sub> and/or osmolality on the growth and recombinant tPA production of CHO cells. *Biotechnol. Bioeng.* 52, 152–160. doi:10.1002/(SICI)1097-0290(19961005)52:1<152::AID-BIT15>3.0.CO;2-Q
- Klengel, T., Liang, W. J., Chaloupka, J., Ruoff, C., Schröppel, K., Naglik, J. R., et al. (2005). Fungal adenyl cyclase integrates CO<sub>2</sub> sensing with cAMP signaling and virulence. *Curr. Biol.* 15, 2021–2026. doi:10.1016/j.cub.2005.10.040
- Klinger, C., Schmidberger, C., and Eisenkraetzer, D. (2015). *Scale Down Strategies of Cell Culture Processes. Oral Presentation at DECHEMA Conference, Scale-Up and Scale-Down of Bioprocesses*. Hamburg.
- Knoll, A., Bartsch, S., Husemann, B., Engel, P., Schroer, K., Ribeiro, B., et al. (2007). High cell density cultivation of recombinant yeasts and bacteria under non-pressurized and pressurized conditions in stirred tank bioreactors. *J. Biotechnol.* 132, 167–179. doi:10.1016/j.jbiotec.2007.06.010
- Knoll, A., Maier, B., Tscherrig, H., and Büchs, J. (2005). The oxygen mass transfer, carbon dioxide inhibition, heat removal, and the energy and cost efficiencies of high pressure fermentation. *Adv. Biochem. Eng. Biotechnol.* 92, 77–99. doi:10.1007/b98918
- Kumar, N., Gammell, P., and Clynes, M. (2007). Proliferation control strategies to improve productivity and survival during CHO based production culture: a summary of recent methods employed and the effects of proliferation control in product secreting CHO cell lines. *Cytotechnology* 53, 33–46. doi:10.1007/s10616-007-9047-6
- Kuriyama, H., Mahakarnchanakul, W., Matsui, S., and Kobayashi, H. (1993). The effects of pCO<sub>2</sub> on yeast growth and metabolism under continuous fermentation. *Biotechnol. Lett.* 15, 189–194. doi:10.1007/BF00133022
- Kusian, B., Sültemeyer, D., and Bowien, B. (2002). Carbonic anhydrase is essential for growth of *Ralstonia eutropha* at ambient CO<sub>2</sub> concentrations. *J. Bacteriol.* 184, 5018–5026. doi:10.1128/JB.184.18.5018-5026.2002
- Li, H., Oppenorth, P. H., Wernick, D. G., Rogers, S., Wu, T.-Y., Higashide, W., et al. (2012). Integrated electromicrobial conversion of CO<sub>2</sub> to higher alcohols. *Science* 335, 1596. doi:10.1126/science.1217643
- Litsanov, B., Brocker, M., and Bott, M. (2012). Towards homosuccinate fermentation: metabolic engineering of *Corynebacterium glutamicum* for anaerobic succinate production from glucose and formate. *Appl. Environ. Microbiol.* 78, 3325–3337. doi:10.1128/AEM.07790-11
- Lopes, M., Belo, I., and Mota, M. (2014). Over-pressurized bioreactors: application to microbial cell cultures. *Biotechnol. Prog.* 30, 767–775. doi:10.1002/btpr.1917
- Lu, S., Eiteman, M. A., and Altman, E. (2009). Effect of CO<sub>2</sub> on succinate production in dual-phase *Escherichia coli* fermentations. *J. Biotechnol.* 143, 213–223. doi:10.1016/j.jbiotec.2009.07.012
- McIntyre, M., and McNeil, B. (1998). Morphogenetic and biochemical effects of dissolved carbon dioxide on filamentous fungi in submerged cultivation. *Appl. Microbiol. Biotechnol.* 50, 291–298. doi:10.1007/s002530051293
- Merlin, C., Masters, M., McAteer, S., and Coulson, A. (2003). Why is carbonic anhydrase essential to *Escherichia coli*? *J. Bacteriol.* 185, 6415–6424. doi:10.1128/JB.185.21.6415-6424.2003



- Mitsuhashi, S., Ohnishi, J., Hayashi, M., and Ikeda, M. (2004). A gene homologous to beta-type carbonic anhydrase is essential for the growth of *Corynebacterium glutamicum* under atmospheric conditions. *Appl. Microbiol. Biotechnol.* 63, 592–601. doi:10.1007/s00253-003-1402-8
- Miziorko, H. M., and Lorimer, G. H. (1983). Ribulose-1, 5-bisphosphate carboxylase-oxygenase. *Annu. Rev. Biochem.* 52, 507–532. doi:10.1146/annurev.bi.52.070183.002451
- Nikolau, B. J., Ohlrogge, J. B., and Wurtele, E. S. (2003). Plant biotin-containing carboxylases. *Arch. Biochem. Biophys.* 414, 211–222. doi:10.1016/S0003-9861(03)00156-5
- Noorman, H. J., Luijckx, G. C. A., Luyben, K. C., and Heijnen, J. J. (1992). Modeling and experimental validation of carbon dioxide evolution in alkalophilic cultures. *Biotechnol. Bioeng.* 39, 1069–1079. doi:10.1002/bit.260391102
- Ohnishi, J., Mitsuhashi, S., Hayashi, M., Ando, S., Yokoi, H., Ochiai, K., et al. (2002). A novel methodology employing *Corynebacterium glutamicum* genome information to generate a new L-lysine-producing mutant. *Appl. Microbiol. Biotechnol.* 58, 217–223. doi:10.1007/s00253-001-0883-6
- Okino, S., Inui, M., and Yukawa, H. (2005). Production of organic acids by *Corynebacterium glutamicum* under oxygen deprivation. *Appl. Microbiol. Biotechnol.* 68, 475–480. doi:10.1007/s00253-005-1900-y
- Okino, S., Noburyu, R., Suda, M., Jojima, T., Inui, M., and Yukawa, H. (2008). An efficient succinic acid production process in a metabolically engineered *Corynebacterium glutamicum* strain. *Appl. Microbiol. Biotechnol.* 81, 459–464. doi:10.1007/s00253-008-1668-y
- Orlowski, A., Vargas, L. A., Aiello, A., and Álvarez, B. V. (2013). Elevated carbon dioxide upregulates NBCn1 Na<sup>+</sup>/HCO<sub>3</sub><sup>-</sup> cotransporter in human embryonic kidney cells. *Am. J. Physiol. Renal Physiol.* 305, F1765–F1774. doi:10.1152/ajprenal.00096.2013
- Ozturk, S. S. (1996). Engineering challenges in high density cell culture systems. *Cytotechnology* 22, 3–16. doi:10.1007/BF00353919
- Ozturk, S. S., and Palsson, B. O. (1991). Effect of medium osmolarity on hybridoma growth, metabolism, and antibody production. *Biotechnol. Bioeng.* 37, 989–993. doi:10.1002/bit.260371015
- Park, S. A., Ko, A., and Lee, N. G. (2011). Stimulation of growth of the human gastric pathogen *Helicobacter pylori* by atmospheric level of oxygen under high carbon dioxide tension. *BMC Microbiol.* 11:96. doi:10.1186/1471-2180-11-96
- Peters-Wendisch, P. G., Schiel, B., Wendisch, V. F., Katsoulidis, E., Möckel, B., Sahm, H., et al. (2001). Pyruvate carboxylase is a major bottleneck for glutamate and lysine production by *Corynebacterium glutamicum*. *J. Mol. Microbiol. Biotechnol.* 3, 295–300.
- Pfizenmaier, J., and Takors, R. (2015). “Host mammalian cell,” in *Industrial Biotechnology*, eds C. Wittmann, and J. Liao (Weinheim: Wiley-VCH Verlag).
- Pronk, J. T., Yde Steensma, H., and Van Dijken, J. P. (1996). Pyruvate metabolism in *Saccharomyces cerevisiae*. *Yeast* 12, 1607–1633. doi:10.1002/(SICI)1097-0061(199612)12:16<1607:AID-YEA70>3.0.CO;2-4
- Radoš, D. X. I., Turner, D. L., Fonseca, L. L., Carvalho, A. L., Blombach, B., Eikmanns, B., et al. (2014). The effect of CO<sub>2</sub> on anaerobic succinate production by *Corynebacterium glutamicum*: carbon flux analysis by <sup>13</sup>C-NMR. *Appl. Environ. Microbiol.* 80, 3015–3024. doi:10.1128/AEM.04189-13
- Repaske, R., Repaske, A. C., and Mayer, R. D. (1974). Carbon dioxide control of lag period and growth of *Streptococcus sanguis*. *J. Bacteriol.* 117, 652–659.
- Richard, L., Guillouet, S. E., and Uribelarrea, J.-L. (2014). Quantification of the transient and long-term response of *Saccharomyces cerevisiae* to carbon dioxide stresses of various intensities. *Process Biochem.* 49, 1808–1818. doi:10.1016/j.procbio.2014.07.020
- Riedel, C., Rittmann, D., Dangel, P., Möckel, B., Petersen, S., Sahm, H., et al. (2001). Characterization of the phosphoenolpyruvate carboxykinase gene from *Corynebacterium glutamicum* and significance of the enzyme for growth and amino acid production. *J. Mol. Microbiol. Biotechnol.* 3, 573–583.
- Ritchie, R. J., Nadolny, C., and Larkum, A. (1996). Driving forces for bicarbonate transport in the cyanobacterium *Synechococcus R-2* (PCC 7942). *Plant Physiol.* 112, 1573–1584.
- Sauer, U., and Eikmanns, B. J. (2005). The PEP-pyruvate-oxaloacetate node as the switch point for carbon flux distribution in bacteria. *FEMS Microbiol. Rev.* 29, 765–794. doi:10.1016/j.femsre.2004.11.002
- Sechenov, M. (1889). Über die Konstitution der Salzlösungen auf Grund ihres Verhaltens zu Kohlensäure. *Z. Phys. Chem.* 4, 117–125.
- Smith, K. S., and Ferry, J. G. (2000). Prokaryotic carbonic anhydrases. *FEMS Microbiol. Rev.* 24, 335–366. doi:10.1016/S0168-6445(00)00030-9
- Spilimbergo, S., and Bertucco, A. (2003). Non-thermal bacteria inactivation with dense CO<sub>2</sub>. *Biotechnol. Bioeng.* 84, 627–638. doi:10.1002/bit.10783
- Stretton, S., and Goodman, A. E. (1998). Carbon dioxide as a regulator of gene expression in microorganisms. *Antonie Van Leeuwenhoek* 73, 79–85. doi:10.1023/A:1000610225458
- Stumm, W., and Morgan, J. J. (1995). *Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters*, 3rd Edn. New York, NY: Wiley.
- Takors, R. (2012). Scale-up of microbial processes: impacts, tools and open questions. *J. Biotechnol.* 160, 3–9. doi:10.1016/j.jbiotec.2011.12.010
- Takuma, S., Hirashima, C., and Piret, J. M. (2007). Dependence on glucose limitation of the pCO<sub>2</sub> influences on CHO cell growth, metabolism and IgG production. *Biotechnol. Bioeng.* 97, 1479–1488. doi:10.1002/bit.21376
- Talley, R. S., and Baugh, C. L. (1975). Effects of bicarbonate on growth of *Neisseria gonorrhoeae*: replacement of gaseous CO<sub>2</sub> atmosphere. *Appl. Microbiol.* 29, 469–471.
- Tashian, R. E. (1989). The carbonic anhydrases: widening perspectives on their evolution, expression and function. *Bioessays* 10, 186–192. doi:10.1002/bies.950100603
- Tripp, B. C., Smith, K., and Ferry, J. G. (2001). Carbonic anhydrase: new insights for an ancient Enzyme. *J. Biol. Chem.* 276, 48615–48618. doi:10.1074/jbc.R100045200
- Wahlund, T. M., Conway, T., and Tabita, F. R. (1996). Bioconversion of CO<sub>2</sub> to ethanol and other compounds. *ACS Div. Fuel Chem. Prepr.* 41, 1403–1405.
- Wieschalka, S., Blombach, B., Bott, M., and Eikmanns, B. J. (2012). Bio-based production of organic acids with *Corynebacterium glutamicum*. *Microb. Biotechnol.* 6, 87–102. doi:10.1111/1751-7915.12013
- Yang, J., Dogovski, C., Hocking, D., Tauschek, M., Perugini, M., and Robins-Browne, R. M. (2009). Bicarbonate-mediated stimulation of RegA, the global virulence regulator from *Citrobacter rodentium*. *J. Mol. Biol.* 394, 591–599. doi:10.1016/j.jmb.2009.10.033
- Zanghi, J. A., Schmelzer, A. E., Mendoza, T. P., Knop, R. H., and Miller, W. M. (1999). Bicarbonate concentration and osmolality are key determinants in the inhibition of CHO cell polysialylation under elevated pCO<sub>2</sub> or pH. *Biotechnol. Bioeng.* 65, 182–191. doi:10.1002/(SICI)1097-0290(19991020)65:2<182:AID-BIT8>3.0.CO;2-D
- Zelle, R. M., de Hulster, E., Kloezen, W., Pronk, J. T., and van Maris, A. J. A. (2010). Key process conditions for production of C<sub>4</sub> dicarboxylic acids in bioreactor batch cultures of an engineered *Saccharomyces cerevisiae* strain. *Appl. Environ. Microbiol.* 76, 744–750. doi:10.1128/AEM.02396-09
- Zhang, X., Wang, X., Shanmugam, K. T., and Ingram, L. O. (2010). L-Malate production by metabolically engineered *Escherichia coli*. *Appl. Environ. Microbiol.* 77, 427–434. doi:10.1128/AEM.01971-10
- Zhu, M. M., Goyal, A., Rank, D. L., Gupta, S. K., Vanden Boom, T., and Lee, S. S. (2005). Effects of elevated pCO<sub>2</sub> and osmolality on growth of CHO cells and production of antibody-fusion protein B1: a case study. *Biotechnol. Prog.* 21, 70–77. doi:10.1021/bp049815s
- Zippin, J. H., Levin, L. R., and Buck, J. (2001). CO<sub>2</sub>/HCO<sub>3</sub><sup>-</sup> responsive soluble adenylyl cyclase as a putative metabolic sensor. *Trends Endocrinol. Metab.* 12, 366–370. doi:10.1016/S1043-2760(01)00454-4

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Blombach and Takors. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Production of cinnamic and *p*-hydroxycinnamic acids in engineered microbes

Alejandra Vargas-Tah and Guillermo Gosset\*

Departamento de Ingeniería Celular y Biotecnología, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Mexico

## OPEN ACCESS

### Edited by:

Alvaro R. Lara,  
Universidad Autónoma  
Metropolitana-Cuajimalpa, Mexico

### Reviewed by:

Akihiko Kondo,  
Kobe University, Japan  
Judith Becker,  
Saarland University, Germany

### \*Correspondence:

Guillermo Gosset,  
Departamento de Ingeniería  
Celular y Biotecnología,  
Instituto de Biotecnología,  
Universidad Nacional Autónoma  
de México, Apdo. Postal 510-3,  
Cuernavaca, Morelos 62210, México  
gosset@ibt.unam.mx

### Specialty section:

This article was submitted to Systems  
Biology, a section of the journal  
Frontiers in Bioengineering and  
Biotechnology

**Received:** 22 May 2015

**Accepted:** 30 July 2015

**Published:** 20 August 2015

### Citation:

Vargas-Tah A and Gosset G (2015)  
Production of cinnamic and  
*p*-hydroxycinnamic acids in  
engineered microbes.  
Front. Bioeng. Biotechnol. 3:116.  
doi: 10.3389/fbioe.2015.00116

The aromatic compounds cinnamic and *p*-hydroxycinnamic acids (pHCAs) are phenylpropanoids having applications as precursors for the synthesis of thermoplastics, flavoring, cosmetic, and health products. These two aromatic acids can be obtained by chemical synthesis or extraction from plant tissues. However, both manufacturing processes have shortcomings, such as the generation of toxic subproducts or a low concentration in plant material. Alternative production methods are being developed to enable the biotechnological production of cinnamic and (pHCAs) by genetically engineering various microbial hosts, including *Escherichia coli*, *Saccharomyces cerevisiae*, *Pseudomonas putida*, and *Streptomyces lividans*. The natural capacity to synthesize these aromatic acids is not existent in these microbial species. Therefore, genetic modification have been performed that include the heterologous expression of genes encoding phenylalanine ammonia-lyase and tyrosine ammonia-lyase activities, which catalyze the conversion of L-phenylalanine (L-Phe) and L-tyrosine (L-Tyr) to cinnamic acid and (pHCA), respectively. Additional host modifications include the metabolic engineering to increase carbon flow from central metabolism to the L-Phe or L-Tyr biosynthetic pathways. These strategies include the expression of feedback insensitive mutant versions of enzymes from the aromatic pathways, as well as genetic modifications to central carbon metabolism to increase biosynthetic availability of precursors phosphoenolpyruvate and erythrose-4-phosphate. These efforts have been complemented with strain optimization for the utilization of raw material, including various simple carbon sources, as well as sugar polymers and sugar mixtures derived from plant biomass. A systems biology approach to production strains characterization has been limited so far and should yield important data for future strain improvement.

**Keywords:** cinnamic acid, *p*-hydroxycinnamic acid, aromatics, metabolic engineering, phenylpropanoids, natural products, biotechnology

## Introduction

Bacteria and plants have the natural capacity for synthesizing a large number of aromatic compounds from simple carbon sources. The shikimate or common aromatic pathway is the main central metabolic branch leading to several biosynthetic pathways that produce various aromatic metabolites (**Figure 1**). The aromatic amino acids L-phenylalanine (L-Phe), L-tyrosine (L-Tyr), and L-tryptophan (L-Trp) are primary metabolites synthesized from simple carbon sources by

plants and bacteria. Secondary metabolites, such as phenylpropanoids, are derived from L-Phe and L-Tyr and are produced mainly by plants. The phenylpropanoid acids cinnamic acid (CA) and *p*-hydroxycinnamic acid (pHCA), also known as coumaric acid, are two metabolites having nutraceutical and pharmaceutical properties (Chemler and Koffas, 2008). They also have applications as precursors of chemical compounds and materials, such as high-performance thermoplastics (Kaneko et al., 2006; Sariaslani, 2007).

Both CA and pHCA are present in plant tissues at a low concentration. Therefore, complex procedures must be employed for their extraction and the yields are usually low. For these reasons, alternative production schemes are being explored. Several microbial species currently employed in biotechnological processes, possess part of the pathways required for CA and pHCA synthesis from simple carbon sources. Thus, by applying genetic engineering techniques to various microbial species, it has been possible to develop production strains with the novel capacity for synthesizing phenylpropanoid acids (Nijkamp et al., 2007; Vannelli et al., 2007a; Limem et al., 2008).

In this review, we focus on recent studies related to the application of genetic engineering strategies for the development of strains derived from *Escherichia coli*, *Pseudomonas putida*, *Streptomyces lividans*, and *Saccharomyces cerevisiae* for the production of the phenylpropanoid acids CA and pHCA. Production process development issues, such as product toxicity and carbon source utilization, are also discussed.

## The Shikimate Pathway and Derived Aromatic Biosynthetic Pathways

The common aromatic pathway or shikimate pathway includes common reactions leading to the specific biosynthetic pathways for L-Phe, L-Tyr, and L-Trp. Most of the reactions of the shikimate aromatic pathway are conserved among bacteria and plants; however, they can differ in terms of specific pathway regulation. The first reaction in the shikimate pathways is the condensation of central carbon metabolism intermediates phosphoenolpyruvate (PEP) and erythrose-4-phosphate (E4P) to yield 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (DAHP) (Figure 1). This reaction is catalyzed by the enzyme DAHP synthase. In bacteria, this step is usually regulated at the enzyme activity level by a feedback-inhibition allosteric mechanism on DAHP synthase. After six enzyme reactions, DAHP is converted to the intermediate chorismate (CHO). This compound is the metabolic branching point where pathways specific for the synthesis of L-Phe, L-Tyr, and L-Trp originate (Maeda and Dudareva, 2012).

The biosynthetic pathway for L-Tyr biosynthesis starts with the conversion of CHO to prephenate (PPA) by the enzyme chorismate mutase (CM). The intermediate PPA is converted to 4-hydroxyphenylpyruvate (HPP) in a reaction catalyzed by prephenate dehydrogenase (PDH). Finally, HPP is transformed to L-Tyr by transamination. Biosynthesis of L-Phe also starts with the conversion of CHO to PPA followed by a reaction yielding phenylpyruvate (PPY) in a reaction catalyzed by the enzyme

prephenate dehydratase (PDT). It is common in bacteria to find bifunctional enzymes where the CM domain is fused to either the PDH or the PDT domain (Patel et al., 1977).

In most bacteria, L-Phe, L-Tyr, and L-Trp are the final products of their biosynthetic pathways. However, in plants and some bacteria, these amino acids are intermediates in pathways for the synthesis of secondary metabolites. The phenylpropanoids are a class of secondary metabolites produced by plants mainly as protection against biotic stress. These compounds also have medicinal use, such as antioxidants, UV screens, anticancer, antiviral, anti-inflammatory, anti-nitric oxide production, and antibacterial agents (Korkina et al., 2011). The first step in the phenylpropanoid pathway is the deamination of L-Phe to generate CA in a reaction catalyzed by the enzyme phenylalanine ammonia lyase (PAL). This compound is then transformed to pHCA by enzyme cinnamate 4-hydroxylase (C4H) (Achnine et al., 2004). pHCA is a precursor for a large number of metabolites including flavonoids and lignans. These compounds have important structural and protective functions in plants (Emiliani et al., 2009). Several of the characterized PAL enzymes can also employ L-Tyr as a substrate, thus displaying tyrosine ammonia lyase (TAL) activity and producing pHCA directly from L-Tyr (Cochrane et al., 2004; Cui et al., 2014). Therefore, the term PAL/TAL is usually employed in naming this type of enzymes. At present, known TAL enzymes have substrate specificity toward both L-Phe and L-Tyr. The specificity to L-Tyr of the TAL from *Rhodobacter sphaeroides* has been modified to L-Phe by changing amino acid residue His89 that is conserved in enzymes with TAL activity to Phe. This result suggests that it might be possible to modify the substrate specificity of enzymes that display PAL activity by employing protein engineering approaches (Louie et al., 2006).

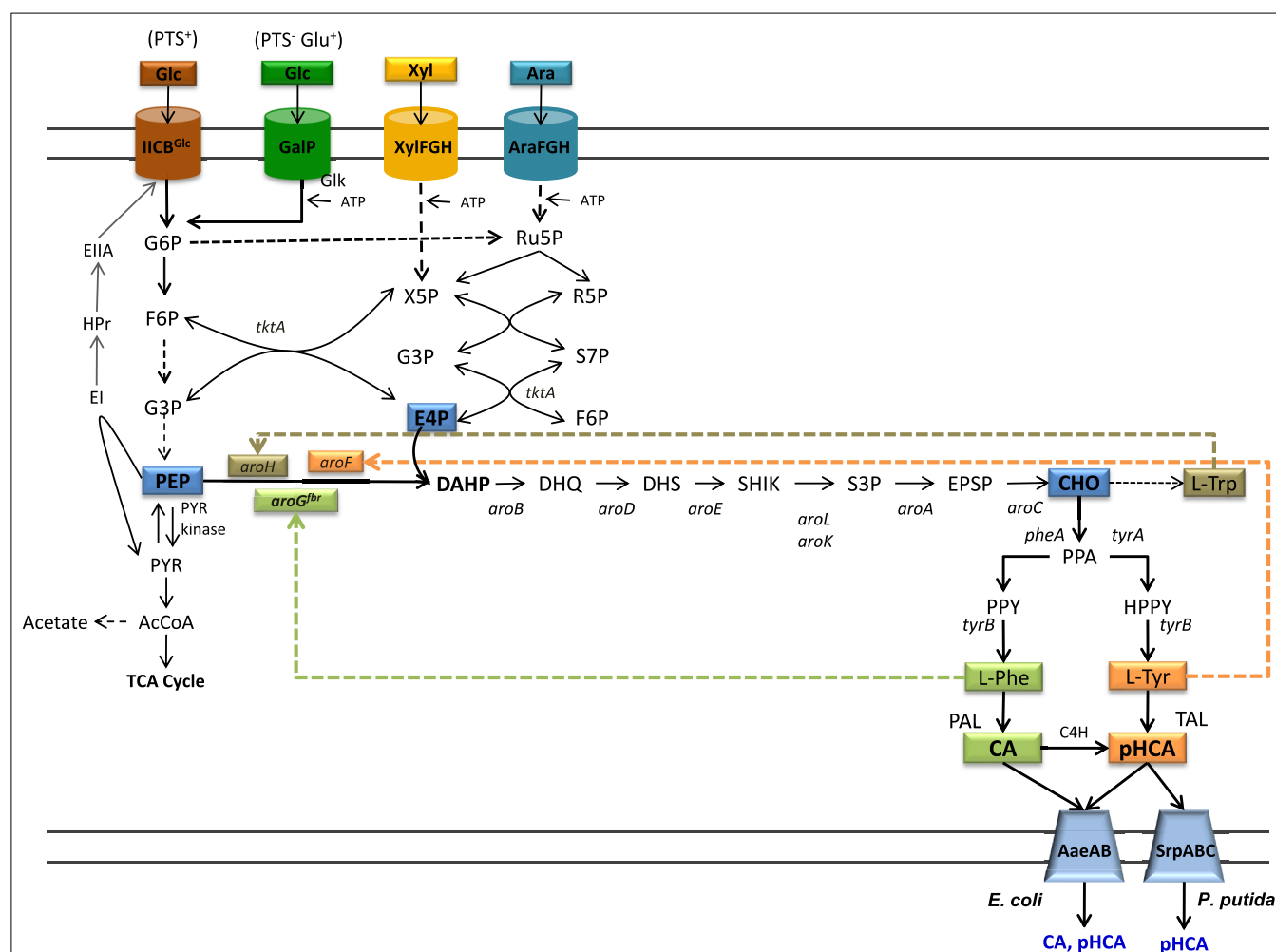
## Engineering of Microbes for Production of Phenylpropanoid Acids

Microbial species traditionally employed in biotechnological processes do not have the natural capacity for synthesizing phenylpropanoid acids. Therefore, genetic modifications are required to generate CA or pHCA production strains. Since enzymes having PAL/TAL activities are not present in industrial microbial strains, a key modification to generate phenylpropanoid acids production strains is the heterologous expression of genes encoding these proteins. These strains are usually further engineered by modifying carbon flow distribution in central and biosynthetic pathways with the aim of increasing L-Phe and L-Tyr synthesis capacity. Such strategies can be complemented with the engineering of substrate utilization and product tolerance. These strategies are presented and discussed in the following sections.

## Engineering of Key Pathways/Targets for the Production of CA and pHCA

### PAL/TAL Enzymes as Key Reactions

The heterologous expression of genes encoding PAL/TAL activities is an essential modification to generate CA and pHCA production strains. Genes encoding PAL/TAL enzymes have been



**FIGURE 1 | Central metabolism, aromatics biosynthetic pathways, and transport pathways from engineered *E. coli*.** Dashed arrows indicate multiple enzyme reactions. EI, PTS enzyme I; HPr, PTS phosphohistidine carrier protein; EIIA, PTS glucose-specific enzyme II; PTS IICB<sup>Glc</sup>, integral membrane glucose permease; GalP, galactose permease; XylFGH, xylose transport proteins; AraFGH, arabinose transport proteins; DAHPs, DAHP synthase; *aroG<sup>fbr</sup>*, gene encoding a feedback-inhibition-resistant version of DAHPs; *tktA*,

transketolase; *tyrB*, tyrosine aminotransferase gene; PAL, phenylalanine ammonia lyase; TAL, tyrosine ammonia lyase; C4H, cinnamate 4-hydroxylase; AaeXAB, efflux pump from *E. coli*; SrpABC, efflux pump from *P. putida*; G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; G3P, glyceraldehyde-3-phosphate; PEP, phosphoenolpyruvate; R5P, ribose-5-phosphate; Ru5P, ribulose-5-phosphate; S7P, sedoheptulose-7-phosphate; X5P, xylulose-5-phosphate; PYR, pyruvate; AcCoA, acetyl-CoA; TCA, tricarboxylic acids.

expressed in various microbial hosts to enable the capacity of transforming L-Phe and L-Tyr into CA and pHCA, respectively. **Table 1** shows the sources of various PAL/TAL enzymes and their kinetic parameters (when known). As it can be observed, the genes employed originate from various biological groups, including bacteria, yeasts, and plants. The kinetic characterization of these enzymes provides data that can be used to compare them regarding substrate affinity and catalytic efficiency. Several enzymes of this family display both PAL and TAL activities. However, a wide range of  $K_m$  values can be observed for substrates L-Phe and L-Tyr (**Table 1**). A microbial strain that expresses a gene encoding a PAL/TAL enzyme acquires the capacity for transforming L-Phe or L-Tyr into the corresponding phenylpropanoid acid. Therefore, in a production context, the aromatic amino acid must be supplemented to the culture medium, where it is then internalized and deaminated by the PAL/TAL enzyme.

## Engineering of Pathway Regulation

Most microbial hosts have the metabolic capacity for synthesizing L-Phe or L-Tyr from simple carbon sources. Therefore, to reduce production costs, it is desirable to enhance endogenous biosynthesis of the aromatic substrates to avoid having to add them to the culture medium. The strategies for generating L-Phe or L-Tyr overproducer strains are well-known and they have been applied to generate microbial strains that can produce these amino acids at grams level from simple carbon sources (Ikeda and Katsumata, 1992; Ikeda et al., 2006; Báez-Viveros et al., 2007; Lütke-Eversloh and Stephanopoulos, 2007, 2008; Chávez-Bejar et al., 2008; Juminaga et al., 2012; Kang et al., 2012). The elimination of enzyme feedback inhibition regulation and transcriptional regulatory processes are the most common modification that enhances flux to the L-Phe or L-Tyr biosynthetic pathway. The high-level expression of feedback inhibition resistant (fbr) versions of enzyme



**TABLE 1 | Kinetic parameters for phenylalanine ammonia lyase and tyrosine ammonia lyase enzymes from various organisms.**

Organism	$K_m$ L-Phe ( $\mu$ M)	$K_m$ L-Tyr ( $\mu$ M)	Reference
<i>Rhodobacter capsulatus</i>	1277	15.6	Kyndt et al. (2002)
<i>Zea mays</i>	270	19	Rosler et al. (1997)
<i>Petroselinum crispum</i>	24.5	7.8	Appert et al. (1994)
<i>Rhodobacter sphaeroides</i>	560	60	Vannelli et al. (2007b)
<i>Phanerochaete chrysosporium</i>	161	44	Xue et al. (2007)
<i>Trichosporon cutaneum</i>	2330	432	Vannelli et al. (2007a)
<i>Rhodotorula minuta</i>	584	212	Vannelli et al. (2007a)
<i>Rhodotorula graminis</i>	448	154	Vannelli et al. (2007a)
<i>Streptomyces maritimus (encP)</i>	23	–	Xiang and Moore (2005)
<i>Rhodotorula glutinis</i>	1340	560	Zhu et al. (2013)
<i>Arabidopsis thaliana (PAL1)</i>	68	–	Cochrane et al. (2004)
<i>Arabidopsis thaliana (PAL2)</i>	64	–	Cochrane et al. (2004)
<i>Arabidopsis thaliana (PAL3)</i>	256	–	Cochrane et al. (2004)
<i>Arabidopsis thaliana (PAL4)</i>	71	–	Cochrane et al. (2004)
<i>Saccharothrix espanaensis</i>	2860	15.5	Berner et al. (2006)

DAHP synthase causes an increase in carbon flow from central metabolism to the shikimate pathway. This increased flux toward CHO synthesis can be redirected to the specific L-Phe and L-Tyr biosynthetic pathway by overexpressing fbr versions of PDT and PDH, respectively. In addition to the aforementioned strategies for increasing synthesis capacity of aromatic amino acids, other approaches, such as a combination of strain random mutagenesis and selection, have proven successful for generating mutants that overproduce L-Phe or L-Tyr (Bongaerts et al., 2001). Random mutagenesis and selection schemes have also been employed to generate *P. putida* strains for the production of CA and pHCA. The *P. putida* strain S12 was isolated in cultures containing a high styrene concentration and has been shown to be solvent tolerant (Weber et al., 1993). This strain was engineered by expressing the gene coding for the PAL/TAL from *R. toruloides*. To increase L-Phe biosynthesis capacity, this strain was subjected to random mutagenesis and a selection process on the toxic analog m-fluoro-phenylalanine. An isolated mutant produced 5 mM CA from glucose (Nijkamp et al., 2005). In addition to CA, this strain produced a very low amount of pHCA. To obtain a high-level pHCA producer strain, random mutagenesis and m-fluoro-phenylalanine selection were employed, considering that L-Phe and L-Tyr share steps in their biosynthetic pathways. Following this procedure, a strain was obtained that showed a 14-fold increase in pHCA synthesis. However, degradation of pHCA was observed in the culture. It is known that *P. putida* has a pCHA catabolic pathway. The gene *fcs* encoding feruloyl-CoA synthetase was inactivated in *P. putida* S12, thus eliminating the first step in the degradative pathway. This modification caused a 2.5-fold increase in pHCA titer (224  $\mu$ M); however, a large amount of CA was also produced (350  $\mu$ M). To reduce CA production, L-Phe auxotrophic mutants were generated by random mutagenesis. One of such mutants produced 860 and 70  $\mu$ M of pHCA and CA, respectively (Nijkamp et al., 2007). These results demonstrate how random mutagenesis schemes coupled to selection with toxic analogs can be employed to successfully yield overproducing strains. However, a drawback of such methods is the lack of knowledge on the

specific mutations responsible for the observed phenotype. The application of genome sequencing to characterize such mutants should yield information that will allow for the future rational design of production strains.

## Engineering Building Blocks Supply

The capacity for synthesizing aromatic amino acids can be improved by increasing availability of central metabolism precursors PEP and E4P. The enzymes transketolase (Tkt) and transaldolase (Tal) from the pentose phosphate pathway participate in E4P metabolism. The overexpression of each enzyme showed a positive effect on DAHP synthesis from glucose. It was found that *tktA* overexpression had a larger positive effect on increasing E4P availability for aromatics biosynthesis. Unexpectedly, simultaneous expression of Tkt and Tal genes did not show a synergistic effect on DAHP synthesis from glucose (Draths et al., 1992; Lu and Liao, 1997). PEP participates in several anabolic and catabolic pathways. In addition, it serves as phosphate donor during import and phosphorylation of glucose and other sugars by the PEP:sugar phosphotransferase system (PTS) (Erni, 2012). When a bacterium grows on a sugar source that can be internalized by the PTS, such as glucose, this system is the major consumer of PEP as one mole of this molecule is required to transport and phosphorylate 1 mole of glucose, generating glucose-6-phosphate (G6P), and PYR (Figure 1). Various studies have stabilized that PEP availability determines the yield of aromatic amino acids synthesized from glucose (Patnaik et al., 1995). The maximal theoretical yield for the synthesis of aromatics precursor DAHP from glucose in a strain with an active PTS is 0.43 mol/mol. However, if glucose phosphorylation is PEP independent, the maximal theoretical yield could be increased twofold to 0.86 mol/mol (Báez et al., 2001). The reactions catalyzed by enzymes Ppc and Pyk consume PEP and for this reason they have become targets for strain improvement. The inactivation of gene *ppc* caused a 10-fold increase in L-Phe production, but  $\mu$  was severely affected (Miller et al., 1987). The enzyme Pps, encoded by *ppsA*, catalyzes a gluconeogenic reaction synthesizing PEP from PYR. The overexpression of *ppsA* in *E. coli* has been shown to increase DAHP production to near the maximum theoretical yield. However, it was also determined that *ppsA* overexpression caused partial growth inhibition (Patnaik et al., 1995).

## Engineering of Substrate Utilization

To increase PEP availability for aromatics production, one approach involves inactivation of PTS activity and its replacement by alternate import and phosphorylation mechanisms. Mutant strains of *E. coli* lacking PTS activity (PTS<sup>−</sup>) have been generated and characterized. These PTS<sup>−</sup> mutants display very low rates of glucose consumption and growth (PTS<sup>−</sup> glucose<sup>−</sup> phenotype). Therefore, they are not useful for production applications. For this reason, several strategies have been followed to improve glucose import capacity in these mutant strains. Starting from a PTS<sup>−</sup> glucose<sup>−</sup> strain, a continuous culture system was employed to select evolved bacterial clones displaying a fourfold higher specific growth rate ( $\mu$ ) (PTS<sup>−</sup> glucose<sup>+</sup> phenotype) (Flores et al., 1996). The characterization of these laboratory evolved strains revealed that glucose uptake and phosphorylation are carried out

by a PTS-independent mechanism involving galactose permease (GalP) and glucokinase (Glk) (Flores et al., 1996, 2002). The enzyme Glk phosphorylates the cytoplasmic glucose employing ATP as the phosphate donor, so in these strains, PEP is not consumed for this reaction. An alternative method for generating PTS<sup>-</sup> glucose<sup>+</sup> strains involves the overexpression of genes encoding native or heterologous proteins having glucose import and ATP-dependent phosphorylating activities (Snoep et al., 1994; Hernández-Montalvo and Martínez, 2003). The characterization of several PTS<sup>-</sup> glucose<sup>+</sup> strains shows that aromatics yield from glucose can be increased to a level close to the maximal theoretical yield calculated for a strain with PEP-independent glucose import (Flores et al., 1996; Báez et al., 2001).

The use of substrates derived from lignocellulosic hydrolyzates is a current trend in the development of biotechnological processes. These are abundant and relatively inexpensive carbon sources that can become the basis of sustainable production processes. Although the composition of such materials differ according to their origin, most of them contain a mixture of pentoses and hexoses, mainly glucose, arabinose, and xylose. The efficient utilization of such sugar mixture by the production strains is an important characteristic. Therefore, microbial species that naturally consume such mixtures must be employed, or genetic modification must be employed to provide such trait. The PTS is involved in carbon catabolic repression, a regulatory process responsible for the sequential utilization of mixtures of carbon sources in *E. coli* and other bacteria. To determine the effect of PTS inactivation on sugar mixtures utilization and aromatics acids production, a combinatorial study was reported where the effect of various phenotypes on CA and pHCA production were compared. The authors generated strains derived from wild type and a PTS<sup>-</sup> glucose<sup>+</sup> mutant that expressed PAL/TAL from *Rhodotorula glutinis* or *A. thaliana* as well as genes encoding an fbr version DAHP synthase and Tkt. These strains were grown in medium supplemented with glucose, arabinose, xylose, or a simulated lignocellulosic hydrolyzate containing a mixture of these three sugars and acetate. When grown in the simulated lignocellulosic hydrolyzate, sequential sugar utilization was observed in the wild-type strain, whereas they were simultaneously consumed by the PTS<sup>-</sup> glucose<sup>+</sup> strain. This is a trait that might prove to increase productivity when employing hydrolyzates from lignocellulosic raw materials (Vargas-Tah et al., 2015).

The organism *S. lividans* has the natural capacity to grow employing various complex carbon sources, including cello-oligosaccharide and xylo-oligosaccharide (Noda et al., 2012). This is a useful trait since no previous physical or chemical treatment of the lignocellulosic biomass is required to generate free sugars. This bacterium was modified for CA production by the heterologous expression of gene *encP* coding for a PAL from *Streptomyces maritimus*. Production of CA was observed when culturing in complex medium supplemented with glucose or glycerol as carbon source. Carbon sources that can be derived from biomass were also tested, including xylose, xylan, and raw starch with CA titers of 300, 130, and 460 mg/L, respectively (Noda et al., 2011). In another report, the *S. lividans* strain expressing *encP* was employed for producing CA from oligosaccharides as carbon sources in complex

medium. This strain produced 490, 400, and 160 mg/L of CA with cello-oligosaccharide, xylo-oligosaccharide, and Avicel as carbon sources, respectively (Noda et al., 2012). A strain of *S. lividans* was constructed for pHCA production by expressing a gene encoding a TAL from *Rhodobacter sphaeroides*. This strain produced 786 and 736 mg/L of pHCA from glucose or cellobiose as carbon source. This strain was further modified by expressing a gene encoding an endoglucanase from *Thermobifida fusca* YX. The recombinant *S. lividans* strain produced 500 mg/L of pHCA from phosphoric acid swollen cellulose (Kawai et al., 2013).

## Engineering of Product Tolerance

It is known that CA and pHCA are toxic compounds for several microorganisms (Qi et al., 2007; Sariaslani, 2007; Vargas-Tah et al., 2015). For example, a concentration of 10 g/L of pHCA completely abolishes growth in *E. coli* (Sariaslani, 2007). Under production conditions where these compounds accumulate to high titers, a negative effect on productivity and cell viability would be expected. Two general approaches have been followed to mitigate the negative effects of CA and pHCA accumulation on production strain performance. The first one is based on the overexpression of efflux systems that can employ pHCA as substrate. In one study, it was demonstrated that *E. coli* mutants in TolC, the outer membrane factor for several efflux systems, are more sensitive to the negative effects of pHCA. The main multidrug efflux system in *E. coli* is AcrAB, it requires TolC to export several kinds of toxic compounds. An increase in sensitivity to pHCA of *acrAB* mutants showed that this compound is a substrate of this efflux system. However, the sensitivity was higher in a strain with mutated efflux system and TolC, suggesting that other(s) efflux system(s) could be active with pHCA. As a strategy to identify gene candidates encoding proteins involved in pHCA efflux, transcriptome analysis was performed with *E. coli* grown in the presence of this aromatic acid. This study found genes *aaeA* and *aaeB* to be upregulated. Genes *aaeXAB* encode an efflux pump, when this operon was overexpressed in *E. coli*, a twofold increase in tolerance to pHCA was observed (Figure 1) (Van Dyk et al., 2004; Sariaslani, 2007). This study also found that *AaeXAB* is functional in the absence of TolC, thus indicating the existence of an additional aromatics acids efflux system that is dependent on TolC (Van Dyk et al., 2004).

The second approach to avoid the toxic effects or aromatic acid is based on the use as production host of an organism having a natural high tolerance to these compounds. The *P. putida* strain S12 was isolated from cultures containing a high concentration of styrene. Characterization of this strain has revealed that it is tolerant to various organic solvents (Weber et al., 1993). An efflux pump encoded by genes *srpABC* in this strain has been identified as an important factor in solvent tolerance and in its capacity to export chemical products (Figure 1). Their expression in a solvent-sensitive *P. putida* strain increases its resistance to solvents (Kieboom et al., 1998). As mentioned above, engineering of this strain has led to production hosts displaying the highest reported CA and pHCA titers in production cultures (Nijkamp et al., 2005, 2007).

## Microbial Hosts Employed for Phenylpropanoid Acids Production

As mentioned above, various microbial strains have been modified for the production of CA or pHCA by employing diverse metabolic engineering strategies. This includes Gram-negative, Gram-positive, and eukaryotic organisms. The Gram-negative bacterium *E. coli* is a facultative anaerobe that can employ a large variety of organic compounds as carbon and energy sources. *E. coli* was the first organism modified by genetic engineering and it is currently employed as an important model in metabolic engineering experiments. The existence of a wide array of genetic modification techniques developed for *E. coli* has enabled the engineering of this organism for aromatic acids production. As shown in **Table 2**, CA and pHCA *E. coli* production strains have been generated that express PAL/TAL enzymes from various origins, having the capacity to employ single sugars or mixtures as substrates. Even though *E. coli* is a useful host organism, increasing its tolerance to aromatic acids is still a challenge for developing robust production strains.

The genus *Streptomyces* includes Gram-positive organisms displaying the natural capacity for synthesizing antibiotics and other secondary metabolites. Among them, the bacterium *S. lividans* synthesizes enzymes that enable it to consume various biomass-derived polymers, such as xylan, starch, cello-oligosaccharide, and xylo-oligosaccharide (Noda et al., 2012). This organism has been genetically modified for production of CA and pHCA from simple sugars or polymers as carbon sources. The rate of production for

the aromatic acids was lower with sugar polymers when compared to simple sugars. This indicates that further strain improvement would be required to increase the cellular activities related to polymer consumption (**Table 2**) (Noda et al., 2011, 2012; Kawai et al., 2013).

The Gram-negative bacterium *P. putida* is a versatile organism, displaying the capacity to colonize diverse niches. A *P. putida* strain designated S12 was isolated from cultures containing a high styrene concentration and it has been shown to be tolerant to several organic solvents (Weber et al., 1993). *P. putida* S12 has been engineered to generate CA and pHCA production strains. These strains have reached the highest aromatic acids titers in production cultures a result likely attributed to the solvent tolerance of the progenitor strain (**Table 2**). The introduction of functions related to sugar polymer consumption would bring *P. putida* S12 closer to an ideal aromatics acids production strain.

The yeast *S. cerevisiae* is a unicellular eukaryotic organism that has been employed as an important biological model. In addition, yeast is an industrial organism that has been fundamental to the development of various fermentative processes. Currently, there is one example of the genetic modification of *S. cerevisiae* for pHCA production. A production strain was generated by expressing the *R. glutinis* PAL/TAL. This *S. cerevisiae* strain also expressed the Cytochrome P-450 enzyme system from the plant *Helianthus tuberosus*. When culturing this yeast strain with glucose or raffinose as carbon sources, the highest titers of pHCA were 14.6 and 202.8 (Vannelli et al., 2007a,b).

**TABLE 2 | Comparison of production parameters for aromatic acids synthesized by engineered microbial strains.**

Organism	PAL/TAL	Carbon source	pHCA (mg/L)	CA (mg/L)	Reference
<i>E. coli</i> (w3110)	<i>R. glutinis</i>	Glucose	7.0	4.9	Vannelli et al. (2007a)
<i>E. coli</i> (DH10B)	<i>R. glutinis</i>	Glucose	46.0	41.5	Vannelli et al. (2007a)
<i>E. coli</i> (ATCC 31884)	<i>R. glutinis</i>	Glucose	24.6	56.3	Vannelli et al. (2007a)
<i>E. coli</i> (BL21*AI)	<i>T. cutaneum</i>	LB	79	186	Vannelli et al. (2007a)
<i>E. coli</i> (PTS <sup>-</sup> )	<i>R. glutinis</i>	Glucose	15.0	29.0	Vargas-Tah et al. (2015)
<i>E. coli</i> (PTS <sup>-</sup> )	<i>A. thaliana</i>	Glucose	–	78.5	Vargas-Tah et al. (2015)
<i>E. coli</i> (W3110)	<i>A. thaliana</i>	Arabinose	–	151.4	Vargas-Tah et al. (2015)
<i>E. coli</i> (W3110, <i>pheA</i> <sup>-</sup> )	<i>R. glutinis</i>	Simulated hydrolyzate	58.4	15.6	Vargas-Tah et al. (2015)
<i>E. coli</i> (W3110)	<i>A. thaliana</i>	Simulated hydrolyzate	–	55.9	Vargas-Tah et al. (2015)
<i>S. cerevisiae</i>	<i>R. glutinis</i>	Glucose	2.3	–	Vannelli et al. (2007a)
<i>S. cerevisiae</i>	<i>R. glutinis</i>	Raffinose	31.8	–	Vannelli et al. (2007a)
<i>P. putida</i> S12 C1	<i>R. toruloides</i>	Glucose	36.8	46.5	Nijkamp et al. (2007)
<i>P. putida</i> S12 C3	<i>R. toruloides</i>	Glucose	141.2	10.4	Nijkamp et al. (2007)
<i>P. putida</i> S12 C3	<i>R. toruloides</i>	Glucose	1740.1	22.2	Nijkamp et al. (2007)
<i>P. putida</i> S12	<i>R. toruloides</i>	Glucose	–	61.4	Nijkamp et al. (2005)
<i>P. putida</i> S12	<i>R. toruloides</i>	Glucose	–	740.3	Nijkamp et al. (2005)
<i>P. putida</i> S12	<i>R. toruloides</i>	Glycerol	–	799.5	Nijkamp et al. (2005)
<i>S. lividans</i>	<i>R. sphaeroides</i>	Glucose	786	–	Kawai et al. (2013)
<i>S. lividans</i>	<i>R. sphaeroides</i>	Cellulobiose	736	–	Kawai et al. (2013)
<i>S. lividans</i> , endoglucanase expressing	<i>R. sphaeroides</i>	Glucose	753	–	Kawai et al. (2013)
<i>S. lividans</i> , endoglucanase expressing	<i>R. sphaeroides</i>	Cellulobiose	743	–	Kawai et al. (2013)
<i>S. lividans</i> , endoglucanase expressing	<i>R. sphaeroides</i>	Cellulose	500	–	Kawai et al. (2013)
<i>S. lividans</i>	<i>S. maritimus</i>	Glucose	–	80.0	Noda et al. (2012)
<i>S. lividans</i>	<i>S. maritimus</i>	Glucose	–	209.8	Noda et al. (2012)
<i>S. lividans</i>	<i>S. maritimus</i>	Glycerol	–	450	Noda et al. (2011)
<i>S. lividans</i>	<i>S. maritimus</i>	Xylose	–	300.0	Noda et al. (2012)
<i>S. lividans</i>	<i>S. maritimus</i>	Xylan	–	130.0	Noda et al. (2011)
<i>S. lividans</i>	<i>S. maritimus</i>	Raw starch	–	460.0	Noda et al. (2011)
<i>S. lividans</i>	<i>S. maritimus</i>	Cello-oligosaccharide	–	490	Noda et al. (2012)

## Production of CA and pHCA Using Biotransformation

As mentioned above, biological synthesis of CA and pHCA starting from simple or complex sugars has the potential for becoming a relatively inexpensive manufacturing scheme. However, such processes must deal with the issue of product toxicity, which places an upper limit on attainable product titers. A solution to this problem would be to employ a manufacturing alternative based on performing a biotransformation where L-Phe and L-Tyr are employed as substrates for the production of CA and pHCA, respectively. In this scheme, the cell host is required only for the synthesis of the PAL/TAL enzyme, which can be purified for performing the deamination reactions. Production of pHCA was studied by comparing PAL/TAL enzymes from yeast *R. glutinis* (RgTAL) and fungus *Phanerochaete chrysosporium* (PcTAL). The genes coding for these enzymes were expressed in *E. coli* and the protein products were purified. Characterization of purified enzymes showed that PcTAL is thermostable, with a maximal activity at 55–60°C. In experiments with whole cells of *E. coli* expressing PcTAL, 42.2 g/L of pHCA were produced with a specific productivity of 1.11 g/g h (Xue et al., 2007). In another report, the RgTAL was stabilized by encapsulation within polyethyleneimine-mediated biomimetic silica. The free enzyme lost all its activity when exposed 1 h to a temperature of 60°C, whereas encapsulated RgTAL retained 43% of its initial activity (Cui et al., 2015). These results show how a natural version of PAL/TAL or one that has been stabilized can constitute viable alternatives for the synthesis of aromatic acids from L-Phe or L-Tyr substrates.

## Conclusion and Outlook

The development of microbial strains for the production of CA and pHCA from simple carbon sources involves extensive engineering of cellular metabolism. In addition to improving production characteristics, these modifications will likely also cause unexpected alterations to the cell's physiology. System biology approaches offer the opportunity of better understanding the consequences of genetic modifications and the responses to various stress factors during the production stage. The application of omics-based approaches, such as transcriptomics, proteomics, fluxomics, and metabolomics, provides a comprehensive view of the cell's physiology response to various genetic modifications as well as environment factors, such as product toxicity. For the case of microbial strains for the production of CA and pHCA, there is a lack of studies based on omics approaches. However, there are some reports focusing on the study of strains modified for the production of precursors of aromatic acids. In one report, proteome analysis was performed to understand the effect of inactivating a PYR kinase PykF in *E. coli*. Among proteins differentially expressed in the mutant strain, some were related to E4P synthesis and the common aromatic pathway, suggesting a higher capacity for aromatics synthesis (Prabhakar et al., 2007). Transcriptome analysis was performed to compare a PTS<sup>+</sup> and a PTS<sup>−</sup> glucose<sup>+</sup> *E. coli* strain modified for L-Phe production. Among differentially expressed genes, it was found that operon

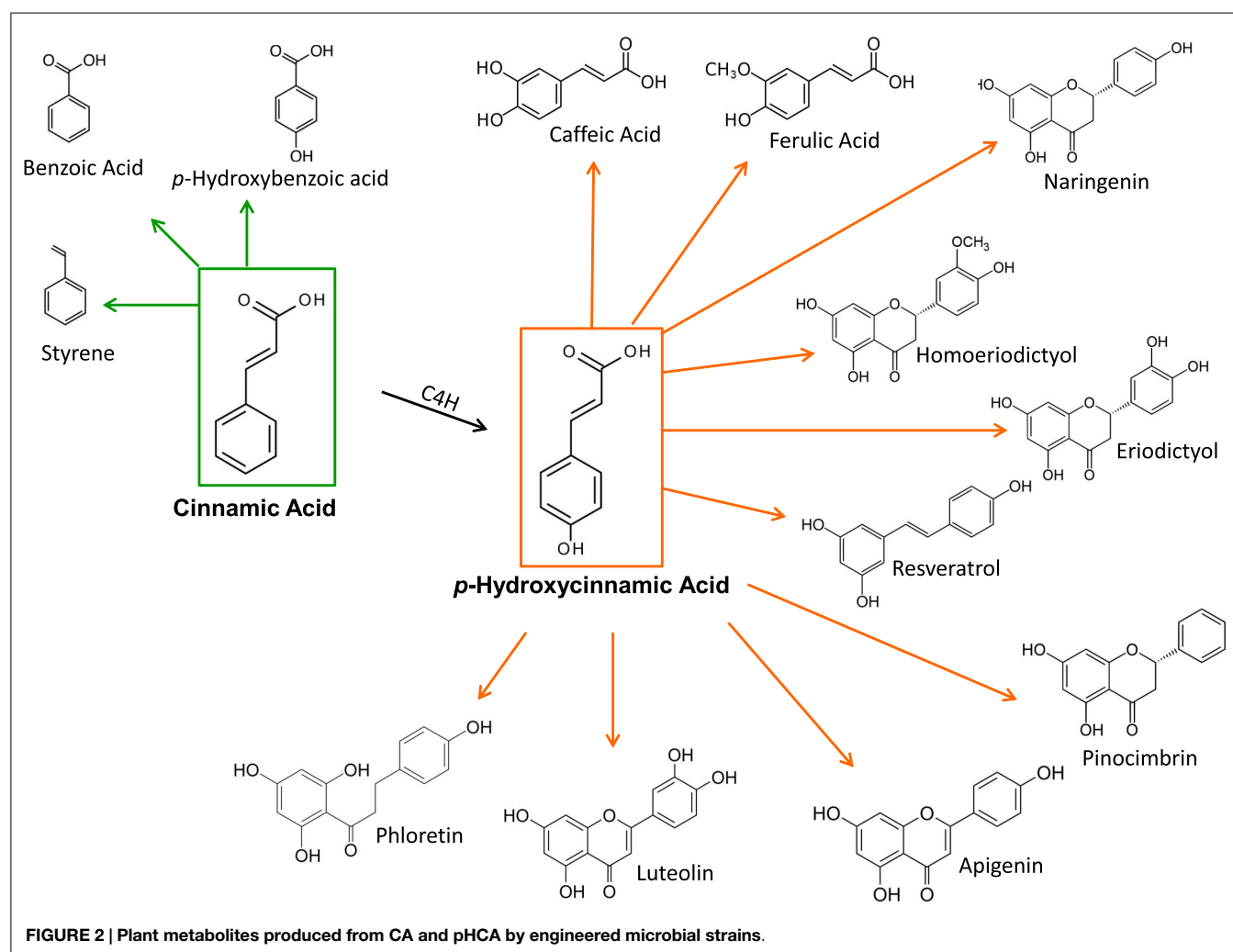
*acs-actP* that is involved in acetate consumption was upregulated in the PTS<sup>−</sup> glucose<sup>+</sup> strain (Báez-Viveros et al., 2007). This response is consistent with the lower level of acetate accumulation in culture medium observed for strain PTS<sup>−</sup> glucose<sup>+</sup> when compared to PTS<sup>+</sup>. These results provide useful data that helps in identifying genetic targets for strain improvement. Future studies focused on characterizing CA and pHCA production strains will likely identify novel targets for strain optimization.

Aromatic acids CA and pHCA are valuable chemicals having direct applications and serving also as precursors for the synthesis of a large number of useful compounds. During the last years, various microbial hosts have been modified by metabolic engineering to generate production strains. These efforts have been fundamental for defining strain development strategies and for identifying factors that limit productivity. In contrast to other biotechnological products where a single microbial host is usually employed, for the case of CA and pHCA production, several different species show promise as production platforms. As reviewed here, *E. coli*, *S. cerevisiae*, *P. putida*, and *S. lividans* display particular characteristics that can favor aromatics acids production. Although much progress has been made with regard to production strain construction and process development, the yields of aromatic acids are still low when compared to other aromatic products (Bongaerts et al., 2001). An important factor limiting productivity is the toxicity of CA and pHCA. In this regard, studies identifying genes encoding efflux systems in *E. coli* and *P. putida* S12 enable a better understanding of the processes involved in mitigating aromatic acids toxicity (Kieboom et al., 1998; Van Dyk et al., 2004). The overexpression of these genes in each organism clearly increases resistance to toxic compounds. It remains to be determined if the solvent-tolerance trait can be transferred to a different species. The use of an omics approach to determine the transcriptional response to CA and pHCA should prove to be valuable for identifying systems that participate in toxic resistance in other microbial species.

Generating a single product is usually the expected outcome in a biotechnological production system. In microbial strains engineered to produce aromatic acids from simple carbon sources, it has been shown that synthesis of CA as only product is possible, as a result of PAK specificity toward L-Phe. However, this is not the case for pHCA, since known TAL enzymes can also employ L-Phe as substrate. Therefore, pHCA is produced always with a certain amount of CA. Although downstream processing could be employed to separate pHCA from CA, this approach would result in increased production costs. Another solution to this issue could be based on applying protein engineering methods to modify substrate specificity of a TAL enzyme for reducing or abolishing CA production, while maintaining high-catalytic activity to increase production of pHCA. As an alternative, the search for novel TAL proteins in natural diversity has the potential for finding enzymes having substrate specificity only toward L-Tyr.

Microbial strains having the capacity for producing CA or pHCA have been employed as platforms for the synthesis of various phenylpropanoid compounds. These include simple phenylpropanoids as well as lignoids, flavonoids, coumarins, and other related compounds (Figure 2) (Dixon and Steele, 1999).





These plant metabolites have been shown to have pharmacological activities, such as antioxidants, anticancer, antiviral, anti-inflammatory, anti-nitric oxide production and antibacterial agents, among others (Dhanalakshmi et al., 2002). The microbial production of these compounds represents an attractive alternative to plant tissue extraction processes. However, at present, these microbial strains produce a low level of these plant compounds. It can be expected that some of the metabolic engineering strategies

applied to CA and pHCA production strains, as reviewed here, should provide a basis for the future improvement of microbial strains that synthesize useful plant metabolites.

## Acknowledgments

This work was supported by CONACyT grant 177568. AV-T was supported by a fellowship from CONACyT.

## References

- Achnine, L., Blancaflor, E. B., Rasmussen, S., and Dixon, R. A. (2004). Colocalization of l-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis. *Plant Cell* 16, 3098–3109. doi:10.1105/tpc.104.024406
- Appert, C., Logemann, E., Hahlbrock, K., Schmid, J., and Amrhein, N. (1994). Structural and catalytic properties of the four phenylalanine ammonia-lyase isoenzymes from Parsley (*Petroselinum crispum* Nym.). *Eur. J. Biochem.* 225, 491–499. doi:10.1111/j.1432-1033.1994.00491.x
- Báez, J. L., Bolívar, F., and Gosset, G. (2001). Determination of 3-deoxy-D-arabino-heptulosonate 7-phosphate productivity and yield from glucose in *Escherichia coli* devoid of the glucose phosphotransferase transport system. *Biotechnol. Bioeng.* 73, 530–535. doi:10.1002/bit.1088
- Báez-Viveros, J. L., Flores, N., Juárez, K., Castillo-España, P., Bolívar, F., and Gosset, G. (2007). Metabolic transcription analysis of engineered *Escherichia coli* strains that overproduce l-phenylalanine. *Microb. Cell Fact.* 6, 30. doi:10.1186/1475-2859-6-30
- Berner, M., Krug, D., Bihlmaier, C., Vente, A., Müller, R., and Bechthold, A. (2006). Genes and enzymes involved in caffeic acid biosynthesis in the actinomycete *Saccharothrix espanaensis*. *J. Bacteriol.* 188, 2666–2673. doi:10.1128/JB.188.7.2666-2673.2006
- Bongaerts, J., Krämer, M., Müller, U., Raeven, L., and Wubbolts, M. (2001). Metabolic engineering for microbial production of aromatic amino acids and derived compounds. *Metab. Eng.* 3, 289–300. doi:10.1006/mben.2001.0196
- Chávez-Béjar, M. I., Lara, A. R., López, H., Hernández-Chávez, G., Martínez, A., Ramírez, O. T., et al. (2008). Metabolic engineering of *Escherichia coli* for L-tyrosine production by expression of genes coding for the chorismate mutase

- domain of the native chorismate mutase-prephenate dehydratase and a cyclohexadienyl dehydrogenase from *Zymomonas mobilis*. *Appl. Environ. Microb.* 74, 3284–3290. doi:10.1128/AEM.02456-07
- Chemler, J. A., and Koffas, M. A. (2008). Metabolic engineering for plant natural product biosynthesis in microbes. *Curr. Opin. Biotechnol.* 19, 597–605. doi:10.1016/j.copbio.2008.10.011
- Cochrane, F. C., Davin, L. B., and Lewis, N. G. (2004). The *Arabidopsis* phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. *Phytochemistry* 65, 1557–1564. doi:10.1016/j.phytochem.2004.05.006
- Cui, J., Liang, L., Han, C., and Lin Liu, R. (2015). Stabilization of phenylalanine ammonia lyase from *Rhodotorula glutinis* by encapsulation in polyethyleneimine-mediated biomimetic silica. *Appl. Biochem. Biotechnol.* 176, 999–1011. doi:10.1007/s12010-015-1624-0
- Cui, J. D., Qiu, J. Q., Fan, X. W., Jia, S. R., and Tan, Z. L. (2014). Biotechnological production and applications of microbial phenylalanine ammonia lyase: a recent review. *Crit. Rev. Biotechnol.* 34, 258–268. doi:10.3109/07388551.2013.791660
- Dhanalakshmi, S., Singh, R. P., Agarwal, C., and Agarwal, R. (2002). Silibinin inhibits constitutive and TNF $\alpha$ -induced activation of NF- $\kappa$ B and sensitizes human prostate carcinoma DU145 cells to TNF $\alpha$ -induced apoptosis. *Oncogene* 21, 1759–1767. doi:10.1038/sj/onc/1205240
- Dixon, R. A., and Steele, C. L. (1999). Flavonoids and isoflavonoids – a gold mine for metabolic engineering. *Trends Plant Sci.* 4, 394–400. doi:10.1016/S1360-1385(99)01471-5
- Draths, K. M., Pompliano, D. L., Conley, D. L., Frost, J. W., Berry, A., Disbrow, G. L., et al. (1992). Biocatalytic synthesis of aromatics from d-glucose: the role of transketolase. *J. Am. Chem. Soc.* 114, 3956–3962. doi:10.1021/ja00036a050
- Emiliani, G., Fondi, M., Fani, R., and Gribaldo, S. (2009). A horizontal gene transfer at the origin of phenylpropanoid metabolism: a key adaptation of plants to land. *Biol. Direct* 4, 7. doi:10.1186/1745-6150-4-7
- Erni, B. (2012). The bacterial phosphoenolpyruvate: sugar phosphotransferase system (PTS): an interface between energy and signal transduction. *J. Iranian Chem. Soc.* 10, 593–630. doi:10.1007/s13738-012-0185-1
- Flores, N., Xiao, J., Berry, A., Bolivar, F., and Valle, F. (1996). Pathway engineering for the production of aromatic compounds in *Escherichia coli*. *Nat. Biotechnol.* 14, 620–623. doi:10.1038/nbt0596-620
- Flores, S., Gosset, G., Flores, N., de Graaf, A. A., and Bolívar, F. (2002). Analysis of carbon metabolism in *Escherichia coli* strains with an inactive phosphotransferase system by <sup>13</sup>C labeling and NMR spectroscopy. *Metab. Eng.* 4, 124–137. doi:10.1006/mben.2001.0209
- Hernández-Montalvo, V., and Martínez, A. (2003). Expression of galP and Glk in a *Escherichia coli* PTS mutant restores glucose transport and increases glycolytic flux to fermentation products. *Biotechnol. Bioeng.* 83, 687–694. doi:10.1002/bit.10702
- Ikeda, M. (2006). Towards bacterial strains overproducing L-tryptophan and other aromatics by metabolic engineering. *Appl. Environ. Microb.* 69, 615–626. doi:10.1007/s00253-005-0252-y
- Ikeda, M., and Katsumata, R. (1992). Metabolic engineering to produce tyrosine or phenylalanine in a tryptophan-producing *Corynebacterium glutamicum* strain. *Appl. Environ. Microb.* 58, 781–785.
- Jumina, D., Baidoo, E. E., Redding-Johanson, A. M., Batth, T. S., Burd, H., Mukhopadhyay, A., et al. (2012). Modular engineering of L-tyrosine production in *Escherichia coli*. *Appl. Environ. Microb.* 78, 89–98. doi:10.1128/AEM.06017-11
- Kaneko, T., Thi, T. H., Shi, D. J., and Akashi, M. (2006). Environmentally degradable, high-performance thermoplastics from phenolic phytomonomers. *Nat. Mater.* 5, 966–970. doi:10.1038/nmat1778
- Kang, S. Y., Choi, O., Lee, J. K., Hwang, B. Y., Uhm, T. B., and Hong, Y. S. (2012). Artificial biosynthesis of phenylpropanoic acids in a tyrosine overproducing *Escherichia coli* strain. *Microb. Cell Fact.* 11, 153. doi:10.1186/1475-2859-11-153
- Kawai, Y., Noda, S., Ogino, C., Takeshima, Y., Okai, N., Tanaka, T., et al. (2013). p-Hydroxycinnamic acid production directly from cellulose using endoglucanase and tyrosine ammonia lyase-expressing *Streptomyces lividans*. *Microb. Cell Fact.* 12, 45. doi:10.1186/1475-2859-12-45
- Kieboom, J., Dennis, J. J., de Bont, J. A., and Zylstra, G. J. (1998). Identification and molecular characterization of an efflux pump involved in *Pseudomonas putida* S12 solvent tolerance. *J. Biol. Chem.* 273, 85–91. doi:10.1074/jbc.273.1.85
- Korkina, L., Kostyuk, V., De Luca, C., and Pastore, S. (2011). Plant phenylpropanoids as emerging anti-inflammatory agents. *Mini Rev. Med. Chem.* 11, 823–835. doi:10.2174/138955711796575489
- Kyndt, J. A., Meyer, T. E., Cusanovich, M. A., and Van Beeumen, J. J. (2002). Characterization of a bacterial tyrosine ammonia lyase, a biosynthetic enzyme for the photoactive yellow protein. *FEBS Lett.* 512, 240–244. doi:10.1016/S0014-5793(02)02272-X
- Limem, I., Emmanuel, G., Alain, H., Frédéric, B., Leila, C. G., Jean-Marc, E., et al. (2008). Production of phenylpropanoid compounds by recombinant microorganisms expressing plant-specific biosynthesis genes. *Process Biochem.* 43, 463–479. doi:10.1016/j.procbio.2008.02.001
- Louie, G. V., Bowman, M. E., Moffitt, M. C., Baiga, T. J., Moore, B. S., and Noel, J. P. (2006). Structural determinants and modulation of substrate specificity in phenylalanine-tyrosine ammonia-lyases. *Chem. Biol.* 13, 1327–1338. doi:10.1016/j.chembiol.2006.11.011
- Lu, J. L., and Liao, J. C. (1997). Metabolic engineering and control analysis for production of aromatics: role of transaldolase. *Biotechnol. Bioeng.* 53, 132–138. doi:10.1002/(SICI)1097-0290(19970120)53:2<132::AID-BIT2>3.0.CO;2-P
- Lütke-Eversloh, T., and Stephanopoulos, G. (2007). L-tyrosine production by deregulated strains of *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 75, 103–110. doi:10.1007/s00253-006-0792-9
- Lütke-Eversloh, T., and Stephanopoulos, G. (2008). Combinatorial pathway analysis for improved L-tyrosine production in *Escherichia coli*: identification of enzymatic bottlenecks by systematic gene overexpression. *Metab. Eng.* 10, 69–77. doi:10.1016/j.ymben.2007.12.001
- Maeda, H., and Dudareva, N. (2012). The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annu. Rev. Plant Biol.* 63, 73–105. doi:10.1146/annurev-arplant-042811-105439
- Miller, J. E., Backman, K. C., O'Connor, M. J., and Hatch, R. T. (1987). Production of phenylalanine and organic acids by phosphoenolpyruvate carboxylase-deficient mutants of *Escherichia coli*. *J. Ind. Microbiol.* 2, 143–149. doi:10.1007/BF01569421
- Nijkamp, K., van Luijk, N., de Bont, J. A., and Wery, J. (2005). The solvent-tolerant *Pseudomonas putida* S12 as host for the production of cinnamic acid from glucose. *Appl. Microbiol. Biotechnol.* 69, 170–177. doi:10.1007/s00253-005-1973-7
- Nijkamp, K., Westerhof, R. G., Ballerstedt, H., de Bont, J. A., and Wery, J. (2007). Optimization of the solvent-tolerant *Pseudomonas putida* S12 as host for the production of p-coumarate from glucose. *Appl. Microbiol. Biotechnol.* 74, 617–624. doi:10.1007/s00253-006-0703-0
- Noda, S., Miyazaki, T., Miyoshi, T., Miyake, M., Okai, N., Tanaka, T., et al. (2011). Cinnamic acid production using *Streptomyces lividans* expressing phenylalanine ammonia lyase. *J. Ind. Microbiol. Biotechnol.* 38, 643–648. doi:10.1007/s10295-011-0955-2
- Noda, S., Miyazaki, T., Miyoshi, T., Miyake, M., Okai, N., Tanaka, T., et al. (2012). Production of *Streptomyces lividans* transglutaminase and cinnamic acid by recombinant *Streptomyces lividans* cultured on biomass-derived carbon sources. *Bioresour. Technol.* 104, 648–651. doi:10.1016/j.biortech.2011.10.045
- Patel, N., Pierson, D. L., and Jensen, R. A. (1977). Dual enzymatic routes to L-tyrosine and L-phenylalanine via pretyrosine in *Pseudomonas aeruginosa*. *J. Biol. Chem.* 252, 5839–5846.
- Patnaik, R., Spitzer, R. G., and Liao, J. C. (1995). Pathway engineering for production of aromatics in *Escherichia coli*: confirmation of stoichiometric analysis by independent modulation of AroG, TktA, and Pps activities. *Biotechnol. Bioeng.* 46, 361–370. doi:10.1002/bit.260460409
- Prabhakar, K., Roshan, C., and Kazuyuki, S. (2007). Proteomic investigation on the Pyk-F gene knockout *Escherichia coli* for aromatic amino acid production. *Enzyme Microb. Technol.* 41, 455–465. doi:10.1016/j.enzmictec.2007.03.018
- Qi, W. W., Vannelli, T., Breinig, S., Ben-Bassat, A., Gatenby, A. A., Haynie, S. L., et al. (2007). Functional expression of prokaryotic and eukaryotic genes in *Escherichia coli* for conversion of glucose to p-hydroxystyrene. *Metab. Eng.* 9, 268–276. doi:10.1016/j.ymben.2007.01.002
- Rosler, J., Krekel, F., Amrhein, N., and Schmid, J. (1997). Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. *Plant Physiol.* 113, 175–179. doi:10.1104/pp.113.1.175
- Sariaslani, F. S. (2007). Development of a combined biological and chemical process for production of industrial aromatics from renewable resources. *Annu. Rev. Microbiol.* 61, 51–69. doi:10.1146/annurev.micro.61.080706.093248
- Snoep, J. L., Arfman, N., Yomano, L. P., Fliege, R. K., Conway, T., and Ingram, L. O. (1994). Reconstruction of glucose uptake and phosphorylation in a glucose-negative mutant of *Escherichia coli* by using *Zymomonas mobilis* genes encoding the glucose facilitator protein and glucokinase. *J. Bacteriol.* 176, 2133–2135.

- Van Dyk, T. K., Templeton, L. J., Cantera, K. A., Sharpe, P. L., and Sariaslani, F. S. (2004). Characterization of the *Escherichia coli* AaeAB efflux pump: a metabolic relief valve? *J. Bacteriol.* 186, 7196–7204. doi:10.1128/JB.186.21.7196-7204.2004
- Vannelli, T., Wei Qi, W., Sweigard, J., Gatenby, A. A., and Sariaslani, F. S. (2007a). Production of p-hydroxycinnamic acid from glucose in *Saccharomyces cerevisiae* and *Escherichia coli* by expression of heterologous genes from plants and fungi. *Metab. Eng.* 9, 142–151. doi:10.1016/j.ymben.2006.11.001
- Vannelli, T., Zhixiong, X., Sabine, B., Wei Qi, W., and Sariaslani, F. S. (2007b). Functional expression in *Escherichia coli* of the tyrosine-inducible tyrosine ammonia-lyase enzyme from yeast *Trichosporon cutaneum* for production of p-hydroxycinnamic acid. *Enzyme Microb. Technol.* 41, 413–422. doi:10.1016/j.enzmictec.2007.03.013
- Vargas-Tah, A., Martínez, L. M., Hernández-Chávez, G., Rocha, M., Martínez, A., Bolívar, F., et al. (2015). Production of cinnamic and p-hydroxycinnamic acid from sugar mixtures with engineered *Escherichia coli*. *Microb. Cell Fact.* 14, 6. doi:10.1186/s12934-014-0185-1
- Weber, F. J., Ooijkaas, L. P., Schemen, R. M., Hartmans, S., and de Bont, J. A. (1993). Adaptation of *Pseudomonas putida* S12 to high concentrations of styrene and other organic solvents. *Appl. Environ. Microbiol.* 59, 3502–3504.
- Xiang, L., and Moore, B. S. (2005). Biochemical characterization of a prokaryotic phenylalanine ammonia lyase. *J. Bacteriol.* 187, 4286–4289. doi:10.1128/JB.187.12.4286-4289.2005
- Xue, Z., McCluskey, M., Cantera, K., Ben-Bassat, A., Sima Sariaslani, F., and Huang, L. (2007). Improved production of p-hydroxycinnamic acid from tyrosine using a novel thermostable phenylalanine/tyrosine ammonia lyase enzyme. *Enzyme Microb. Technol.* 42, 58–64. doi:10.1016/j.enzmictec.2007.07.025
- Zhu, L., Cui, W., Fang, Y., Liu, Y., Gao, X., and Zhou, Z. (2013). Cloning, expression and characterization of phenylalanine ammonia-lyase from *Rhodotorula glutinis*. *Biotechnol. Lett.* 35, 751–756. doi:10.1007/s10529-013-1140-7

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Vargas-Tah and Gosset. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# ColE1-plasmid production in *Escherichia coli*: mathematical simulation and experimental validation

Inga Freudenau<sup>1</sup>, Petra Lutter<sup>1</sup>, Ruth Baier<sup>2</sup>, Martin Schleef<sup>2</sup>, Hanna Bednarz<sup>1,3</sup>,  
Alvaro R. Lara<sup>4</sup> and Karsten Niehaus<sup>1,3\*</sup>

## OPEN ACCESS

### Edited by:

Firas H. Kobeissy,  
University of Florida, USA

### Reviewed by:

Lyamine Hedjazi,  
Institute of Cardiomatobiology and  
Nutrition (ICAN), France  
Zaher Dawy,  
American University of Beirut,  
Lebanon

### \*Correspondence:

Karsten Niehaus,  
Abteilung für Proteom- und  
Metabolomforschung, Fakultät für  
Biologie, Universität Bielefeld,  
Universitätsstr. 25, Bielefeld 33615,  
Germany  
kniehaus@cebitec.uni-bielefeld.de

### Specialty section:

This article was submitted to  
Systems Biology, a section of the  
journal Frontiers in Bioengineering  
and Biotechnology

**Received:** 05 June 2015

**Accepted:** 13 August 2015

**Published:** 01 September 2015

### Citation:

Freudenau I, Lutter P, Baier R,  
Schleef M, Bednarz H, Lara AR and  
Niehaus K (2015) ColE1-plasmid  
production in *Escherichia coli*:  
mathematical simulation and  
experimental validation.  
Front. Bioeng. Biotechnol. 3:127.  
doi: 10.3389/fbioe.2015.00127

<sup>1</sup>Abteilung für Proteom- und Metabolomforschung, Fakultät für Biologie, Universität Bielefeld, Bielefeld, Germany,

<sup>2</sup>PlasmidFactory GmbH & Co. KG, Bielefeld, Germany, <sup>3</sup>Institut für Genomforschung und Systembiologie, Centrum für  
Biotechnologie (CeBiTec), Universität Bielefeld, Bielefeld, Germany, <sup>4</sup>Departamento de Procesos y Tecnología, Universidad  
Autónoma Metropolitana-Cuajimalpa, Mexico City, Mexico

Plasmids have become very important as pharmaceutical gene vectors in the fields of gene therapy and genetic vaccination in the past years. In this study, we present a dynamic model to simulate the ColE1-like plasmid replication control, once for a DH5 $\alpha$ -strain carrying a low copy plasmid (DH5 $\alpha$ -pSUP 201-3) and once for a DH5 $\alpha$ -strain carrying a high copy plasmid (DH5 $\alpha$ -pCMV-lacZ) by using ordinary differential equations and the MATLAB software. The model includes the plasmid replication control by two regulatory RNA molecules (RNAI and RNAII) as well as the replication control by uncharged tRNA molecules. To validate the model, experimental data like RNAI- and RNAII concentration, plasmid copy number (PCN), and growth rate for three different time points in the exponential phase were determined. Depending on the sampled time point, the measured RNAI- and RNAII concentrations for DH5 $\alpha$ -pSUP 201-3 reside between  $6 \pm 0.7$  and  $34 \pm 7$  RNAI molecules per cell and  $0.44 \pm 0.1$  and  $3 \pm 0.9$  RNAII molecules per cell. The determined PCNs averaged between  $46 \pm 26$  and  $48 \pm 30$  plasmids per cell. The experimentally determined data for DH5 $\alpha$ -pCMV-lacZ reside between  $345 \pm 203$  and  $1086 \pm 298$  RNAI molecules per cell and  $22 \pm 2$  and  $75 \pm 10$  RNAII molecules per cell with an averaged PCN of  $1514 \pm 1301$  and  $5806 \pm 4828$  depending on the measured time point. As the model was shown to be consistent with the experimentally determined data, measured at three different time points within the growth of the same strain, we performed predictive simulations concerning the effect of uncharged tRNA molecules on the ColE1-like plasmid replication control. The hypothesis is that these tRNA molecules would have an enhancing effect on the plasmid production. The *in silico* analysis predicts that uncharged tRNA molecules would indeed increase the plasmid DNA production.

**Keywords:** plasmid replication, small RNA, modeling, ordinary differential equations, uncharged tRNA, high copy plasmid, biotechnology



## Introduction

Since the finding that genetically engineered DNA can be used for gene therapy and DNA vaccination in the early 90s last century, the interest in plasmid DNA (pDNA) as a pharmaceutical gene vector has increased constantly (Kutzler and Weiner, 2008). The investigation of DNA vaccination in the past years shows enticing results in several areas, especially in prophylactic vaccine strategies and the usage of pDNA as potential therapeutics for the treatment of infectious diseases, cancer, Alzheimer disease, and allergies (Kutzler and Weiner, 2008; Mairhofer and Lara, 2014). A very promising field of application might be the genetic vaccination with DNA molecules to induce an immune response (Schleef and Blaesens, 2009). Using DNA molecules in that field has the advantage that there are no safety concerns associated with live vaccines. Additionally, the manufacturing process is short and stable in contrast to conventional vaccines (Kutzler and Weiner, 2008; Lara and Ramírez, 2012). It is expected that, since DNA vaccines and non-viral gene therapies enter phase 3 clinical trials and are approved for utilization, the demand for pDNA will increase (Bower and Prather, 2009). To meet these requirements, industrial-scale production processes of pDNA with adequate bacterial strains and vectors have to be designed. Therefore, it is important to understand which factors influence the plasmid replication.

In this work, we present a mathematical model to simulate the regulation of ColE1-like plasmid replication. Bacterial plasmids control their copy number through negative regulatory feedback mechanisms that adjust the replication rate (del Solar and Espinosa, 2000). In the case of ColE1-like plasmids, the replication is controlled by an antisense RNA molecule, which inhibits the maturation of the necessary primer RNA molecule for the DNA polymerase I (Grabherr and Bayer, 2002). When replication occurs, a part of an RNA preprimer transcript, named RNAII, binds to the plasmid origin of replication (*oriV*), where it forms a persistent hybrid (Itoh and Tomizawa, 1980). In the next step, it must be cleaved by the enzyme RNase H, which is specific for DNA–RNA hybrids (Schumann, 2001). This cleavage is essential to release the 3'OH group, so that the elongation by DNA polymerase I can start (Itoh and Tomizawa, 1980). To prevent the replication, this modification of RNAII by RNase H can be blocked through binding of a small complementary RNA transcript, named RNAI, which is the specific inhibitor of primer formation (Tomizawa and Itoh, 1981). RNAI and RNAII are coded within the same DNA region that is part of the *oriV* (Tomizawa, 1984). RNAI is constitutively transcribed from the opposite strand and regulates the frequency of replication initiation (Tomizawa, 1984). RNAI and RNAII form a transient, the so-called “kissing complex,” thus the RNase H cannot modify the RNA preprimer transcript anymore (Tomizawa, 1984, 1985). This complex is stabilized by the binding of the Rom protein (also called Rop protein) (Tomizawa, 1986, 1990). Additionally, the plasmid replication control can be affected by starvation conditions, which leads to a drop in free amino acids, and thus to an increase of uncharged tRNA molecules (Yavachev and Ivanov, 1988). If an uncharged tRNA molecule binds to an RNAI or RNAII molecule, no kissing complex can be built and

the RNAII transcript can serve as a preprimer (Grabherr and Bayer, 2002). Consequently, a high amount of uncharged tRNA molecules is correlated with an increase in plasmid copy number (PCN) (Wróbel and Węgrzyn, 1998). Two models are proposed to describe the interaction between uncharged tRNA molecules and the regulatory RNA molecules. Yavachev and Ivanov (1988) found structural and sequence similarities between the loops of RNAI, respectively, RNAII and the cloverleaf structure of certain tRNA molecules. The second model was introduced by Wang et al. (2002), who suggest that the binding between an uncharged tRNA molecule and one of the regulatory RNA molecules takes place at the amino acid binding site. Although there are kinetic models, which explain the regulation of the ColE1-like plasmid replication (Ataai and Shuler, 1986; Bremer and Lin-Chao, 1986; Keasling and Palsson, 1989a,b; Brendel and Perelson, 1993; Wang et al., 2002), these models consider only parts of the ColE1-like plasmid replication control. Brendel and Perelson (1993), for instance, describe a model for the *in vivo* replication control mechanism by the regulatory RNA molecules, which includes kinetic information and accounts for measured concentration values.

In this study, the model proposed by Brendel and Perelson was extended. Their model describes the ColE1-like plasmid replication control by RNAI and RNAII molecules, with or without the interaction of Rom protein. They investigated plasmid concentrations at two different growth rates and showed a decrease of the PCN in presence of an intact *rom* gene as well as an increase of plasmid production at low growth rates (Brendel and Perelson, 1993). Apart from the replication control by the two regulatory RNA molecules, the plasmid replication can be influenced by uncharged tRNA molecules under amino acid starvation conditions (Wróbel and Węgrzyn, 1998). The mathematical model proposed in the present study incorporates the ColE1-like plasmid replication control by RNAI and RNAII molecules with or without the Rom protein. Additionally, the regulation by uncharged tRNA molecules is described. The advantage of this model is that it is confirmed by *in vitro* measurements of the plasmid concentration at different time points and the appropriate growth rates. Additionally, the free intracellular RNAI and RNAII concentrations were determined via quantitative reverse transcription real-time-PCR (qRT-PCR) for the same time points. With these data, the model was fitted and validated, so it could be used for *in silico* analysis of the ColE1-like plasmid replication control. Since this model takes into account the regulation by modified tRNA molecules, which cannot be charged by amino-acyl-tRNA synthetases anymore, it is possible to investigate the effect on the ColE1-like plasmid replication control.

## Materials and Methods

### Bacterial Strain and Plasmids

The *Escherichia coli* strain DH5 $\alpha$  (F- $\Phi$ 80*lacZ* $\Delta$ M15  $\Delta$ (*lacZ*YA-*argF*) U169 *recA1 endA1 hsdR17* (rK $^-$ , mK $^+$ ) *phoA supE44*  $\lambda$ -*thi-1 gyrA96 relA1*) (Source: Plasmid Factory, Bielefeld, Germany) was used as a host strain for transformation of the high copy plasmid pCMV-*lacZ* (Source: Plasmid Factory, Bielefeld, Germany) as well as the low copy plasmid pSUP 201-3

(Simon et al., 1983). pCMV-lacZ is a ColE1-derived high copy plasmid for therapeutic pDNA production with a size of 7164 bp. It contains a pBR322 origin of replication without the *rom* gene and carries genetic elements for DNA vaccination together with the ampicillin-resistance gene *bla*. The low copy plasmid pSUP 201-3 is also a ColE1-derivative and has a size of 7896 bp. It contains a pBR322 origin of replication together with the *rom* gene and carries a specific recognition site for mobilization. For selection, the ampicillin-resistance gene *bla* and a chloramphenicol resistance gene are located on pSUP 201-3.

## Cultivation

The cultivation started under aerobic conditions with a pre-culture in a shaker at 37°C and 300 rpm in LB medium (Invitrogen, Darmstadt, Germany). Then, the cells were transferred to a synthetic minimal medium (prepared according to Lara et al., 2008). The medium consists of: 5 g·L<sup>-1</sup> Glucose, 17 g·L<sup>-1</sup> K<sub>2</sub>HPO<sub>4</sub>, 5.3 g·L<sup>-1</sup> KH<sub>2</sub>PO<sub>4</sub>, 2.5 g·L<sup>-1</sup> (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1 g·L<sup>-1</sup> NH<sub>4</sub>Cl, 1 g·L<sup>-1</sup> NaCl, 0.01 g·L<sup>-1</sup> Thiamine hydrochloride, 1 g·L<sup>-1</sup> MgSO<sub>4</sub>·7H<sub>2</sub>O, and 1 mL L<sup>-1</sup> trace elements solution. The trace elements solution contains: 17 g·L<sup>-1</sup> Zn(CH<sub>3</sub>COO)<sub>2</sub>·2H<sub>2</sub>O, 7 g·L<sup>-1</sup> Na-EDTA, 1.25 g·L<sup>-1</sup> CoCl<sub>2</sub>·6H<sub>2</sub>O, 7.5 g·L<sup>-1</sup> MnCl<sub>2</sub>·4H<sub>2</sub>O, 0.75 g·L<sup>-1</sup> CuCl<sub>2</sub>·2H<sub>2</sub>O, 1.5 g·L<sup>-1</sup> H<sub>3</sub>BO<sub>3</sub>, 1.05 g·L<sup>-1</sup> Na<sub>2</sub>MoO<sub>4</sub>·2H<sub>2</sub>O, 50 g·L<sup>-1</sup> Fe(III) citrate. All chemicals were purchased from Sigma-Aldrich (Seelze, Germany), Roth (Karlsruhe, Germany), VWR (Darmstadt, Germany), Merck (Darmstadt, Germany), Fluka (Seelze, Germany), and Serva (Heidelberg, Germany). No ampicillin was added in order to avoid antibiotic stress influence on the plasmid production. During the cultivation, samples were successively taken, the cells were harvested by centrifugation (1 mL culture at 10,000 × g), then the supernatant was removed and the pellet was frozen in liquid nitrogen for 10 min and afterwards stored at -80°C. This was done for three different time points during the exponential phase of the cultivation, since we are interested in replication during cultivation. Three biological replicates were taken at each time point. Samples taken during the exponential growth phase are assumed to represent a quasi steady state of the metabolism and plasmid replication and the most representative phase of a pDNA batch production process.

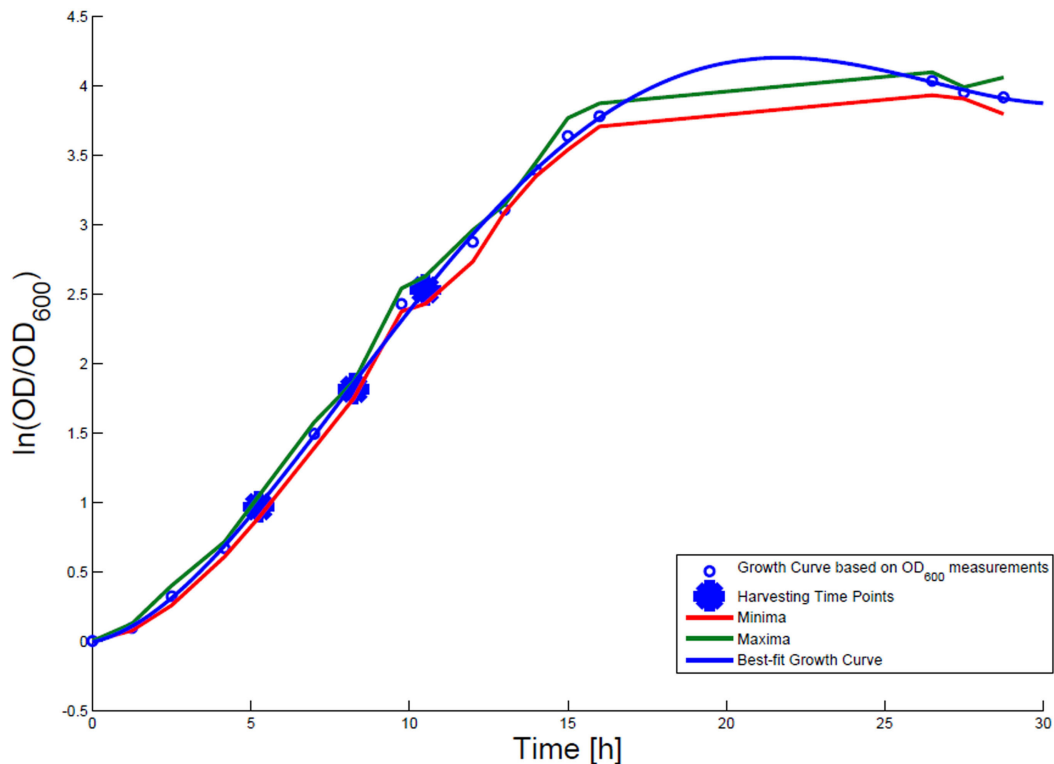
## Growth Rate Determination

For growth rate determination at all measured time points, the cells were cultured in minimal medium and the optical density at 600 nm was measured. Growth curves were generated, where the optical density data as the natural logarithm of the measured OD<sub>600</sub> values are given on the y-axis, normalized to the initial OD<sub>600</sub> value, and the time is given on the x-axis. These growth curves were fitted applying the Matlab function “polyfit.” This function *p* finds the coefficients of a polynomial *p*(*x*) of degree *n* that fits the optical density data stored in a vector *y* best in a least-squares sense, where *p* is a row vector of length *n* + 1 containing the polynomial coefficients in descending powers, *p*(1)\**x*<sup>*n*</sup> + *p*(2)\**x*<sup>*n*-1</sup> + ... + *p*(*n*)\**x* + *p*(*n* + 1). Afterwards a  $\chi^2$ -test was applied to test the quality of the fit. The successfully fitted growth curves are shown in Figures 1 and 2.

In the next step, the first derivative *p*'(*x*), which enables the calculation of the growth rate for every time point, was generated.

## Determination of the RNAI and RNAII Concentrations

To determine the concentrations of RNAI and RNAII, an internal standard for each of them was synthesized. Therefore, a primer with a T7 RNA polymerase promoter region extension, one for the *rnaI* gene and one for the *rnaII* gene (Table 1), was designed and a PCR amplification was used to amplify double strand RNAI- and RNAII-DNA. The PCR mixture with a final volume of 25  $\mu$ L contained 2.5  $\mu$ L GoTaq reaction buffer, 1.5  $\mu$ L deoxynucleotides (0.2 mM final concentration), 0.5  $\mu$ L forward primer and 0.5  $\mu$ L reverse primer (each with a final concentration of 10 pmol  $\mu$ L<sup>-1</sup>), 0.75  $\mu$ L MgCl<sub>2</sub> (final concentration of 1.5 mM), 1  $\mu$ L GoTaq DNA Polymerase, 0.5  $\mu$ L pDNA, and nuclease-free water to fill up to 25  $\mu$ L. All chemicals were obtained from Promega (Mannheim, Germany) except the PCR primers, which were purchased from Metabion (Martinsried, Germany). The PCR was performed in a MJ Research PTC-100 Programmable Thermal Controller from Lab Recyclers (Gaithersburg, USA) applying the following program: an activation step for the hot start Taq polymerase was carried out at 94°C for 2 min, followed by an initial denaturation step at 94°C for 30 s, an annealing step at 64°C for RNAI (67°C for RNAII) and an extension step at 72°C for 20 s. This cycle was repeated 30 times and ended with a final extension step at 72°C for 2 min. The PCR product was purified using the NucleoSpin Extract II Kit from Macherey-Nagel (Düren, Germany) and the concentration was measured spectrophotometrically (NanoDrop2000c spectrometer, Thermo Scientific, Braunschweig, Germany). The PCR product was used as a template for the T7 RNA polymerase (Roche, Mannheim, Germany) according to the manufacturer's instructions to synthesize RNA. The reaction mixture with a final volume of 20  $\mu$ L contained 1 mM of each nucleotide (NTP), 2  $\mu$ L supplied buffer (10 $\times$ ), 40 U T7 RNA polymerase, 20 U Ribolock RNase inhibitor, 1  $\mu$ g template DNA, and DEPC-treated water to fill up to 20  $\mu$ L. The mixture was incubated for 2 h at 37°C. The residual template DNA was digested by adding 2  $\mu$ L DNaseI and incubated 15 min at 37°C. All chemicals were obtained from Fermentas (Schwerte, Germany). The inactivation was done by lithium chloride (LiCl)-precipitation using the following protocol: 2.5  $\mu$ L LiCl (4M) were added to the 20  $\mu$ L reaction mixture; in the next step, 75  $\mu$ L ethanol (>99%) were added. Then, the RNA was precipitated for 30 min at -80°C. After that, for pelleting the RNA, the samples were centrifuged for 15 min at 4°C with 13,000 rpm. The supernatant was discarded and the pellet was washed with 200  $\mu$ L ethanol (70%). After that it was centrifuged for 15 min at 4°C with 13,000 rpm and the supernatant was discarded. The washing procedure was repeated and after discarding the supernatant, the pellet was dried until the residual ethanol was evaporated. In the last step, the pellet was dissolved in 20  $\mu$ L nuclease-free water. For generating a calibration for RNAI, a 10-fold serial dilution series ranging from  $2 \times 10^{-1}$  to  $2 \times 10^{-6}$  ng  $\mu$ L<sup>-1</sup> was made and analyzed via qRT-PCR. A calibration curve was generated by plotting the cycle threshold values (CT-values) against the time. For the RNAII standard curve, the same procedure was applied. Only the range for the measured 10-fold serial dilution series was different:  $2 \times 10^{-1}$  to  $2 \times 10^{-7}$  ng  $\mu$ L<sup>-1</sup>.

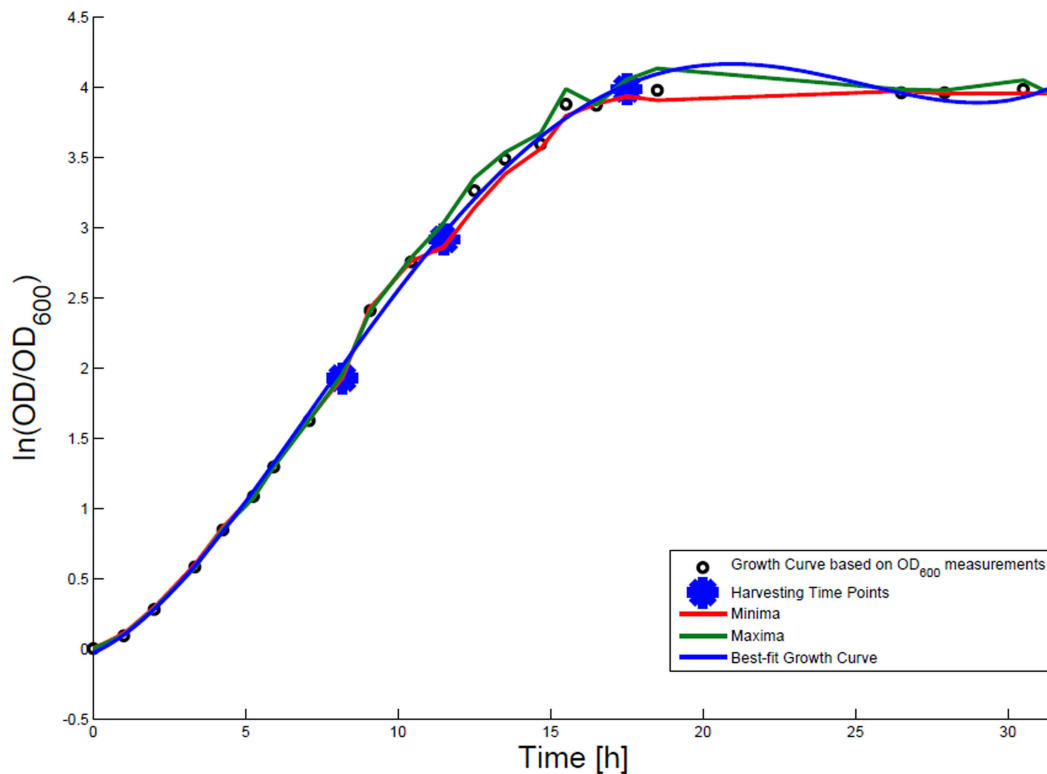


**FIGURE 1 | Growth curve for *E. coli* DH5α-pSUP 201-3.** The growth curve based on OD<sub>600</sub> measurements are marked by black circles and the appropriate best-fit curve is indicated in blue. The harvesting time points  $T_1$ ,  $T_2$ , and  $T_3$  are

presented as bold big points. The minimum–maximum area, within which the measurement values have to reside, is bordered by the red curve (minimal measured OD<sub>600</sub> values) and the green curve (maximal measured OD<sub>600</sub> values).

For sampling the RNA I and RNA II content, cells were cultivated as described above and harvested at each time point. Before harvesting, a dilution series was generated. For different dilutions, 50  $\mu$ L per sample were plated to determine the cell titer. After doing this, 1 mL of the cell culture was harvested and centrifuged. Then, the supernatant was removed and the pellet was frozen in liquid nitrogen for 10 min and then stored at  $-80^\circ\text{C}$ . Lysing Matrix B tubes (MP Biomedicals, Heidelberg, Germany) were prepared with 700  $\mu$ L RLT buffer (RNeasy Plus Mini Kit, Qiagen, Hilden, Germany) with 7  $\mu$ L  $\beta$ -mercaptoethanol. The filled tubes were incubated on ice for 5 min. The pellet was resuspended in 200  $\mu$ L 10 mM Tris-HCl, pH 8, and added to the prepared ice-cold tubes. The cells were lysed using the Hybaid ribolyser for 30 s, level 6.5  $\text{m s}^{-1}$ , and then incubated on ice for 3 min. After centrifugation at 13,000 rpm and  $4^\circ\text{C}$  for 3 min, the supernatant was transferred into a new RNase free reaction tube. The centrifugation step was repeated and again the supernatant was transferred onto the gDNA Eliminator spin column and centrifuged twice at 10,000 rpm,  $4^\circ\text{C}$ , and 30 s. The flow through was transferred into a new RNase free reaction tube. The next steps of RNA isolation were performed using the RNeasy Plus Mini Kit (Qiagen, Hilden Germany) according to the protocol Appendix D, starting at step D3 to step D8, with the modification that after step D6 an additional washing step with 500  $\mu$ L ethanol (80%) was performed. The isolated RNA was measured via quantitative reverse transcription real-time

PCR (qRT-PCR) using the Sensi-Mix SYBR No-Rox One Step Kit (Bioline, Luckenwalde, Germany). The reaction mixture was prepared according to the proposed protocol with the modification that the reverse transcription was done with only one primer (RNAI-reverse and RNAII-reverse). The second primer was added just before the beginning of the PCR. As the RNAI molecule is an antisense RNA to the RNAII molecule, we had to stick to our modified protocol in order to avoid that the respective forward primer would bind to both RNA molecules. In this case, both RNA molecules would be reversely transcribed and we would not be able to determine the absolute amount of the particular molecule by PCR. **Table 2** shows the used primer sets. The reverse transcription reaction was done in the PCR thermo cycler machine (MJ Research PTC-100 Programmable Thermal Controller from Lab Recyclers, Gaithersburg, MD, USA) at  $42^\circ\text{C}$  for 10 min. All real-time-PCR experiments were performed using the Opticon machine (BioRad, München, Germany) according to the following protocol: 10 min at  $95^\circ\text{C}$  initial polymerase activation step, 40 cycles [15 s at  $95^\circ\text{C}$ , annealing 15 s at  $61^\circ\text{C}$  (for RNAI), respectively,  $56^\circ\text{C}$  (for RNAII), 15 s at  $72^\circ\text{C}$  elongation]. A reaction mixture with a final volume of 20  $\mu$ L containing 6.2  $\mu$ L DEPC-treated  $\text{H}_2\text{O}$ , 10.8  $\mu$ L Sensi-Mix, 1  $\mu$ L Primer (RNAI-reverse primer respectively RNAII-reverse primer), and 1  $\mu$ L RNA was used. After having run the reverse transcription, 1  $\mu$ L of the second primer was added and the qRT-PCR was started. After each PCR-cycle, the fluorescence was measured for



**FIGURE 2 | Growth curve for *E. coli* DH5 $\alpha$ -pCMV-lacZ.** The growth curve based on OD<sub>600</sub> measurements is marked by small black circles and the appropriate best-fit curve is indicated in blue. The harvesting time points  $T_1$ ,  $T_2$ , and  $T_3$  are presented as bold big points.

The minimum–maximum area, within which the measurement values have to reside, is bordered by the red curve (minimal measured OD<sub>600</sub> value) and the green curve (maximal measured OD<sub>600</sub> value).

**TABLE 1 | Primer sequences to generate internal RNAI and RNAII standards.**

Target	Orientation	Primer (5'–3')	Length [nt]
<i>mal</i>	Forward	GAAATTAATACGACTCACTATAGG GACAGTATTTGGTATCTGCGCTC	47
	Reverse	AACCACCGCTACCGCGG	18
<i>mall</i>	Forward	GAAATTAATACGACTCACTATAGGG GCAACAAAAAACCACCGCTACCA	50
	Reverse	TTTCATAGGCTCCGCCCC	20

all PCR products. The raw data analysis was carried out with the corresponding Opticon monitor software. Different dilutions of the internal RNA standards were determined via qRT-PCR and the CT-values were calculated. Then the CT-values were plotted against the number of RNA molecules and straight calibration lines for RNAI and RNAII was generated. CT-values were determined as mentioned above for DH5 $\alpha$ -low copy RNA samples and DH5 $\alpha$ -high copy RNA samples and the appropriate absolute RNA concentration was calculated with the help of the linear regression curve. Then, the RNA amount per cell was calculated for each sample in mol per liter. These values were subsequently used for the theoretical model.

**TABLE 2 | Primer sequences to determine the RNAI and RNAII amount via qRT-PCR.**

Target	Orientation	Primer (5'–3')	Length [nt]
<i>mal</i>	Forward	TTGGTATCTGCGCTCTGC	18
	Reverse	CAGCGGTGGTTTGTTC	18
<i>mall</i>	Forward	TAAGTGGCTTCAGCA GAGCGCAGAT	25
	Reverse	CCTGCCGCTTACCG GATACCTGT	23

## Determination of the Plasmid DNA Concentrations

The determination of the pDNA concentrations was done by the company Plasmid Factory GmbH & Co. KG (Bielefeld, Germany). The plasmids were isolated by the Plasmid Factory GmbH & Co. KG (Bielefeld, Germany) with a NucleoBond PC20 kit from Macherey-Nagel (Düren, Germany) and the amount of purified pDNA was measured photometrically at the wavelength of 260 nm. Subsequently, the isolated pDNA was separated by agarose gel electrophoresis to check for possible contamination by chromosomal DNA. In case of contaminations, the ratio of chromosomal DNA to pDNA was determined



densitometrically using the software LabImage 1D L 340 (Intas Science Imaging GmbH). With this ratio, the amount of pure pDNA in the entire isolated pDNA sample could be determined and with the previously determined cell titers the absolute plasmid concentration in PCN for each harvesting time point was calculated. All measurements were carried out with three biological replicates.

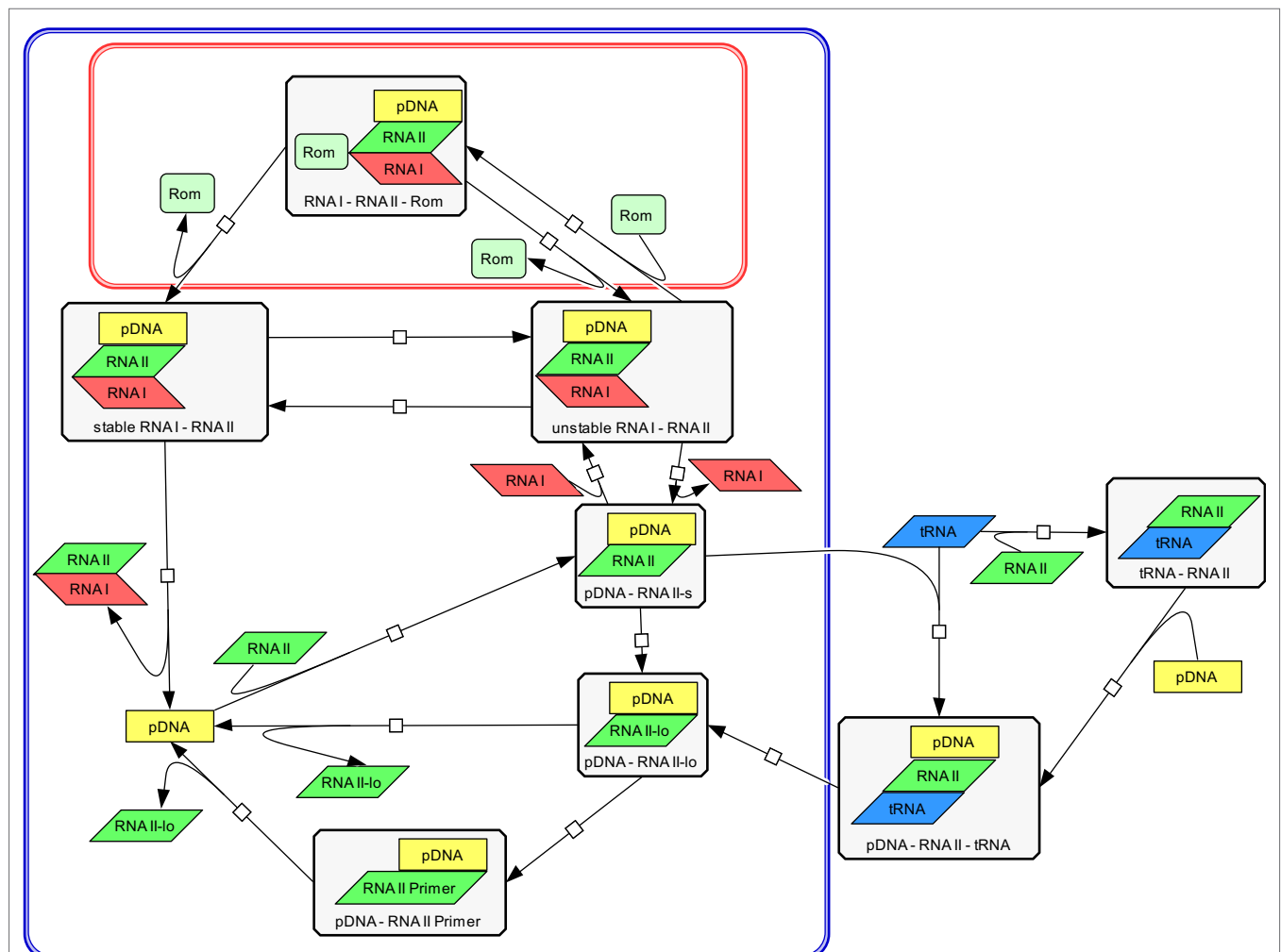
## The Theoretical Model

The modeling work started with building a structural model to visualize all relevant components and their interactions involved in ColE1-like plasmid replication control. This was done using the CellDesigner software (Ver. 4.1) (Funahashi et al., 2008), in which the replication control of low copy and high copy plasmids are described simultaneously (Figure 3). Note that during low copy plasmid control, three additional reactions representing the Rom protein control are involved. These reactions [see rate equations

$v_1$ ,  $v_2$ , and  $v_{13}$  in the ordinary differential equation (ODE) system] are marked with a red box in Figure 3 and describe the regulation by the Rom protein, which does not take place in strains carrying the high copy plasmid, due to deletion of the *rom* gene (Yanisch-Perron et al., 1985; Lin-Chao et al., 1992). As mentioned before, this model is an extension of the model presented by Brendel and Perelson (1993), the reactions which originated from their model are marked with a blue box in Figure 3. The remaining reactions describe the regulation by uncharged tRNA molecules under starvation conditions.

## The Dynamic Model

Based on this structural model, a dynamic model was programed in MATLAB. In that model, each interaction is characterized by a mathematical equation, thereby depicting a kinetic law to define the reaction rates. To describe the dynamics of the substrate concentrations, a system of ODE-System was built,



**FIGURE 3 | Structural model of ColE1-like plasmid replication control for low and high copy plasmids as outlined by the CellDesigner software.** The structural model represents an extension of the model proposed by Brendel and Perelson (1993) for ColE1-like plasmid replication control (the corresponding reactions are surrounded by the blue box). The difference

between the description of the replication control for a low copy plasmid and for a high copy plasmid is three additional reactions (shown in the red box). These reactions describe the control by the Rom protein and are not active in a high copy plasmid. The reactions outside of the blue box were added within this work and describe the replication control by uncharged tRNA molecules.

which can be solved numerically, once the initial conditions are set. For replication control by RNAI and RNAII molecules with and without Rom protein, the rate equations recommended by Brendel and Perelson (1993) were used. As for the regulation by uncharged tRNA molecules, we propose kinetic equations based on mass action law. The corresponding parameters are assumed to be stable over time. Clearly, this assumption would not hold, if all growth phases were to be considered. Changing enzyme concentrations along with varying kinetic behavior would long for different models. As our area of interest is the exponential growth phase, we restrict our model to that time interval, within which the parameters stay constant. *In vivo* measurements of plasmid concentrations together with the appropriate growth rates and the RNAI and RNAII concentrations at different time points were incorporated in the dynamic model as initial conditions. The complete ODE-system along with the rate equations can be found in detail in the Supplementary Material.

## Results

### Cultivation and Determination of the Growth Rate for *E. coli* DH5 $\alpha$ Carrying a Low or a High Copy Plasmid

*Escherichia coli* DH5 $\alpha$ -pSUP 201-3 was cultivated in minimal medium and the optical densities of the culture were determined. These measurements were fitted with the Matlab function “polyfit,” which calculates the coefficient parameter set of polynomial coefficients with the lowest residual sum of squares. That is way, a polynomial equation with an order of four was obtained, which describes the bacterial growth for the strain carrying a low copy plasmid. **Figure 1** presents the appropriate polynomial growth curve fit for *E. coli* DH5 $\alpha$ -pSUP 201-3 together with the measured optical density values and the harvesting time points at which the RNAI, RNAII, and plasmid concentrations were measured. Additionally, the range of the highest and lowest OD<sub>600</sub> values which were measured for each measuring time point is shown (**Figure 1**). In analogy to the low copy plasmid strain, the measured optical densities of the high copy strain were fitted with the MATLAB function “polyfit.” The resulting polynomial equation with an order of four describes the bacterial growth of *E. coli* DH5 $\alpha$ -pCMV-lacZ (**Figure 2**). Comparing the measured optical density values with the best-fit curves (**Figures 1 and 2**), a very good description of the growth behavior for the low copy strain as well as for the high copy strain was obtained. To support this

visual impression,  $\chi^2$ -tests were applied, which returned reduced  $\chi^2$ -values indicating the high fit quality. For both strains, the best-fit curves are located very closely to the measured OD<sub>600</sub> values and fit well into the area bounded by the highest and lowest OD<sub>600</sub> values measured between or at least very close to the minimal and maximal determined optical density curve throughout the whole exponential phase (**Figures 1 and 2**). During this period of cultivation, the theoretical model longs for the respective growth rates: the first derivative of the appropriate polynomial  $p$  gives the growth rate  $\mu$ , for the respective harvesting time point. The calculated growth rates and the appropriate generation times are listed in **Table 3** for *E. coli* DH5 $\alpha$ -pSUP 201-3 and in **Table 4** for DH5 $\alpha$ -pCMV-lacZ.

### Determination of the RNAI/RNAII and the pDNA Concentrations

The RNAI and RNAII concentrations were determined for three different time points by qRT-PCR and the pDNA concentration was measured photometrically via NanoDrop and checked by agarose gel electrophoresis. This was carried out for the low copy strain, *E. coli* DH5 $\alpha$ -pSUP 201-3, and for the high copy strain, *E. coli* DH5 $\alpha$ -pCMV-lacZ, with three biological replicates. After qRT-PCR, the RNA concentrations were calculated by generating straight calibration lines from the RNAI and RNAII standards through linear regression. Regarding the results of *E. coli* DH5 $\alpha$ -pSUP 201-3, a decrease of the RNAI and RNAII concentrations with proceeding growth of the bacterial culture was observed (**Table 5**). Additionally, it was apparent that at every harvesting time point, there are always more free RNAI molecules than RNAII molecules in the cell. In contrast to the RNA concentrations, the measured plasmid concentration per cell was stable over all three harvesting time points. PCNs of 46–48 were calculated for each harvesting time point (**Table 5**). Regarding the determined PCNs of *E. coli* DH5 $\alpha$ -pCMV-lacZ, an increase in plasmid concentrations with proceeding growth of the bacterial culture was observed (**Table 6**). For harvesting time point  $T_1$ , a PCN of 1514 per cell was determined, while at  $T_2$  a PCN of 2403 was found. At  $T_3$ , 5805 plasmids per cell were measured. Cooper and Cass (2004) reported a PCN of 500–700 for pUC derivatives, so compared to the literature, the measured PCNs are very high. For this reason, we double checked the high magnitude of some of these values by qRT-PCR. The PCNs obtained by this analysis were in the same order of magnitude as the previously determined values. Considering the RNA concentrations per

**TABLE 3 |** Calculated growth rates and generation times for DH5 $\alpha$ -pSUP 201-3.

Time point	Harvesting time point [h]	Growth rate [h <sup>-1</sup> ]	Generation time [min]
$T_1$	5.25	0.278	150
$T_2$	8.25	0.302	138
$T_3$	10.5	0.283	147

These parameters are used in the described *in silico* analysis with the dynamic model, when the simulations begin at  $T_1$ ,  $T_2$ , or  $T_3$ . In these simulations, the simulation time corresponds to the generation time and the respective growth rate is used as a parameter in the rate equations.

**TABLE 4 |** Calculated growth rates and generation times for DH5 $\alpha$ -pCMV-lacZ.

Time point	Harvesting time point [h]	Growth rate [h <sup>-1</sup> ]	Generation time [min]
$T_1$	8.17	0.305	136
$T_2$	11.50	0.256	162
$T_3$	17.50	0.090	347

These parameters are used in the described *in silico* analysis with the dynamic model, when the simulations begin at  $T_1$ ,  $T_2$ , or  $T_3$ . In these simulations, the simulation time corresponds to the generation time and the respective growth rate is used as a parameter in the rate equations.

**TABLE 5 | Measured RNAI-, RNAII-, and plasmid concentrations for DH5 $\alpha$ -pSUP at the three harvesting time points, depicted in Figure 1.**

Time point	RNAI [molecules/cell]	RNAII [molecules/cell]	PCN [molecules/cell]
$T_1$	34 ( $\pm 7$ )	3 ( $\pm 0.9$ )	46 ( $\pm 26$ )
$T_2$	17 ( $\pm 5$ )	1 ( $\pm 0.4$ )	48 ( $\pm 30$ )
$T_3$	6 ( $\pm 0.7$ )	0.44 ( $\pm 0.1$ )	46 ( $\pm 17$ )

**TABLE 6 | Measured RNAI-, RNAII-, and plasmid concentrations for DH5 $\alpha$ -pCMV-lacZ at the three harvesting time points, depicted in Figure 2.**

Time point	RNAI [molecules/cell]	RNAII [molecules/cell]	PCN [molecules/cell]
$T_1$	541 ( $\pm 51$ )	22 ( $\pm 2$ )	1514 ( $\pm 1301$ )
$T_2$	1086 ( $\pm 298$ )	64 ( $\pm 4$ )	2403 ( $\pm 713$ )
$T_3$	345 ( $\pm 203$ )	75 ( $\pm 10$ )	5806 ( $\pm 4828$ )

cell, fluctuations in the RNAI concentrations were observed in contrast to measurements for *E. coli* DH5 $\alpha$ -pSUP 201-3, while the RNAII concentrations increased with proceeding growth of the bacterial culture (Table 6). The highest RNAI concentration was measured at  $T_2$  with  $1086 \pm 298$  RNAI molecules per cell and the highest RNAII concentration was found at  $T_3$  with  $75 \pm 10$  molecules per cell. If these RNA concentrations are compared to the RNA concentrations measured for *E. coli* DH5 $\alpha$ -pSUP 201-3, it is apparent that at every harvesting time point, the amount of free RNAI and RNAII molecules in *E. coli* DH5 $\alpha$ -pCMV-lacZ is higher than those in *E. coli* DH5 $\alpha$ -pSUP 201-3.

For determination of the number of RNAI/RNAII molecules per cell as well as the PCN, the mean  $M$  ( $i = 1, 2, 3$ ) for each biological replicate was calculated applying

$$M_i = \frac{1}{N_i} \times \sum_{j=1}^{N_i} m_{ji} \quad (i = 1, 2, 3).$$

The respective SD was determined by

$$SD_i = \sqrt{\frac{1}{N_i} \times \sum_{j=1}^{N_i} (m_{ji} - M_i)^2}.$$

## The Structural Model of ColE1-like Plasmid Replication Control

Using the CellDesigner software (Funahashi et al., 2008), a structural model for the ColE1-like plasmid replication control was built (Figure 3). The involved elements and their interactions are given to describe the ColE1-like plasmid replication control in high- and low copy plasmids simultaneously. Three additional reactions make the difference between low copy plasmid replication control and high copy plasmid replication control. These reactions represent the control by the Rom protein and are marked with a red box in Figure 3. In case of a high copy plasmid replication, the control level by the Rom protein does not take place due to the deletion of the *rom* gene (Yanisch-Perron et al., 1985; Lin-Chao et al., 1992). Since the model proposed in this study is an extension of the model presented by Brendel and Perelson (1993),

the reactions which came from their model are marked by a blue box in Figure 3. The remaining reactions that were added in this study describe the regulation by uncharged tRNA molecules under starvation conditions. Regarding the replication control by uncharged tRNA molecules, there are two possibilities for the uncharged tRNAs (tRNA) to act: the uncharged tRNA molecules can bind to plasmid bound RNAII molecules or they can bind to free RNAII molecules and afterwards build a complex with the pDNA. In both cases, the binding of the inhibitory RNAI molecules is prohibited and the elongation of the RNAII-primer can occur. According to Brendel and Perelson (1993, ) the synthesis and degradation reactions of RNAI, RNAII, and the Rom protein were included. Since the model proposed in this study contains the regulatory level of uncharged tRNA molecules, it was necessary to add also the synthesis and degradation reactions of uncharged tRNA. These synthesis and degradation reactions are not shown in the structural model depicted in Figure 3, they are found in the Supplementary Material.

The plasmid replication depicted in Figure 3 starts with the transcription of a preprimer RNAII molecule from a region 555 bp upstream of the origin of replication. This RNAII transcript has a length of 100–360 nt and persistently hybridizes with a region close to the origin of replication (pDNA–RNAII-s) (Tomizawa, 1986). At that step, there are three possible scenarios for the proceeding of the replication process: the first one is that a regulatory RNA molecule, named RNAI, binds to the hybridized RNAII molecule. The regulatory RNAI molecule is encoded on the opposite strand of that one on which RNAII is encoded. The synthesis of RNAI begins 445 nt upstream of the origin of replication and proceeds in the antisense direction compared to the RNAII synthesis. Since the antisense RNAI transcript is complementary to RNAII, it is able to bind the plasmid–RNAII hybrid and forms a so-called kissing complex (pDNA–RNAII–RNAI unstable). When this transient complex is built, the elongation of RNAII is blocked (Tomizawa, 1986). The second scenario is that the hybridized RNAII molecule is elongated to a length >360 bp before an RNAI molecule could bind (pDNA–RNAII-lo). In that case, RNAII is cleaved by an enzyme named RNase H. This enzyme releases the 3'OH group, so the modified RNAII molecule can act as a primer (pDNA–RNAII-Primer). In the following reaction, the RNAII-Primer will be elongated by the DNA polymerase I and the pDNA will be doubled (Itoh and Tomizawa, 1980). As in the model of Brendel and Perelson (1993), we assume that the concentration and the activity of RNase H remain constant. Additionally, it was assumed that RNAI, RNAII, the Rom protein, and uncharged tRNA molecules move freely through the cytoplasm.

The third scenario is the binding of a free uncharged tRNA to the plasmid bound RNAII (pDNA–RNAII–tRNA). When an uncharged tRNA is bound to RNAII, the RNAI molecules cannot bind anymore, but RNAII can still be cleaved and elongated and serves in this way as a primer (pDNA–RNAII-Primer) (Grabherr and Bayer, 2002). It is proposed in this model that not only plasmid bound RNAII can interact with an uncharged tRNA molecule but also free RNAII molecules can be bound by an uncharged tRNA (RNAII–tRNA). In that case, the RNAII–tRNA complex binds to the template DNA and a complex composed

of a plasmid, RNAII, and an uncharged tRNA molecule is built (pDNA–RNAII–tRNA). The RNAII molecule will be elongated and becomes a primer, so the plasmid replication starts. This third scenario was not considered in the model of Brendel and Perelson. Considering the case that the primer could not be formed because of binding of the RNAI molecule to the plasmid bound RNAII, an unstable complex between the plasmid, RNAII and RNAI is built (pDNA–RNAII–RNAI unstable). This complex could become stable, modeled by a rate equation with the rate constant  $k_{10}$  (Brendel and Perelson, 1993). Having a low copy plasmid, this unstable complex could be stabilized by the Rom protein. As proposed in the model from Brendel and Perelson (1993), we assume that therefore the Rom protein binds to the unstable complex and builds a transient complex (pDNA–RNAII–RNAI–Rom). This transient complex then is converted into a stable complex between of plasmid, RNAII, and RNAI (pDNA–RNAII–RNAI stable). Since the Rom protein is not present in a high copy plasmid, the reactions with Rom participation are missing (Figure 3). In addition to the reactions presented in the structural model in Figure 3, we consider the synthesis and degradation reactions of RNAI, RNAII, uncharged tRNA, and of the Rom protein (Rom).

### The Dynamic Model of ColE1-like Plasmid Replication Control

In the dynamic model, the *in vivo* reaction kinetics of each interaction presented in the structural model (Figure 3) are described by applying mass action law kinetics. In addition to those reactions shown in Figure 3, the reaction rates of the synthesis and degradation reactions of RNAI, RNAII, uncharged tRNA, and Rom protein were specified by mathematical equations. Due to the assumption that the kinetic parameters do not change through the exponential growth phase, they remain constant for every growth rate. The dynamics of the substrate concentrations are described by ODEs. In general, such an ODE reads as follows:

$$\frac{d[S]}{dt} = v_1 + \dots + v_n - \mu \times [S] \quad (1)$$

with  $[S]$  representing the substrate concentration,  $v_i$  ( $i = 1 \dots n$ ) describing the involved reaction rates,  $t$  the time, and  $\mu$  the specific growth rate. In every ODE a term, where the negative growth rate is multiplied with the substrate concentration ( $-\mu \dots [S]$ ), was included. The application of the  $-\mu \dots [S]$  term was proposed by Brendel and Perelson (1993) to account for the dilution effect due to cell growth. The dilution effect describes that the distribution of intracellular compounds to the daughter cell after cell division depends on the growth rate. In case of a fast growing cell, the intracellular volume will increase faster than the concentration of intracellular compounds. As a consequence, the daughter cell shows a lower concentration of intracellular compounds than the mother cell. A lower dilution effect was observed in slow growing cells. The reason for that is that a slow growing mother cell has more time to synthesize her intracellular compounds before it will divide into two daughter cells. Thus, the characteristics of the dilution effect depend on the growth rate. The ODE system is applied to simulate the regulation of ColE1-like plasmid replication. It was set up based on the structural model by applying mass

action law together with the dilution term. The model implies 15 metabolites ( $s_i = 1, \dots, 15$ ) whose interactions are described by 26 rate equations ( $v_i = 1, \dots, 26$ ) containing 26 kinetic constants  $k_i$ . This means that all reactions visualized in the structural model by a reaction arrow are described mathematically by a rate equation ( $v_1, \dots, v_{17}$ ). In addition to the visualized reactions in the structural model, synthesis and degradation reactions of Rom protein, RNAI/RNAII, and uncharged tRNA together with the reaction for complex formation of RNAI and RNAII are described by rate equations ( $v_{18}, \dots, v_{26}$ ). The model proposed in this work is a minimal model which includes the plasmid replication control by two regulatory RNA molecules (RNAI and RNAII), the control by Rom protein as well as the replication control by uncharged tRNA molecules. It comprises no reaction not involved in these levels of replication control. The mentioned 26 rate equations are necessary and sufficient to simulate the ColE1-like plasmid replication control, because they describe all relevant elements to investigate the proposed hypothesis.

The whole ODE-system and the 26 rate equations can be found in the supplement. The parameter values  $k_i$  ( $i = 1, \dots, 26$ ) used in the rate equations with the respective source are listed in Table 7. After defining all kinetic constants, the ODE-system could be solved numerically by applying a MATLAB solver. As initial conditions, the experimentally determined RNAI and RNAII concentrations together with 50% of the respective determined appropriate plasmid concentrations and the appropriate measured growth rates were used. The motivation for taking the half of the measured plasmid concentration value is the assumption that the measured plasmid concentration is equal to the plasmid concentration directly before cell division and that the pDNA will be doubled during one generation time. All unknown substrate concentrations were assigned the zero value, because it is assumed that the internal pool of these substances has to be synthesized and is comparatively low at the beginning of the investigation. Once the above-mentioned parameters and initial concentrations were included into the dynamic model, it was used for *in silico* analysis of the ColE1-like plasmid replication control.

### Simulations of the ColE1-like Plasmid Replication Control – Simulations for Parameter Fitting

Since some kinetic values are unknown so far, they had to be either estimated or determined via parameter fitting. To obtain reasonable fits for the constants  $k_3$ ,  $k_4$ , and  $k_{26}$  (Table 7), the experimentally determined data of harvesting time point  $T_2$  were used for the low copy plasmid as well as for the high copy plasmid. Obviously, this data set was not considered for later validation. In all simulations, one single cell was considered. The simulation time corresponds to the generation time, which was calculated using the respective growth rates determined for each harvesting time point. The simulations for parameter fitting and validation for DH5 $\alpha$ -pSUP 201-3 began at the particular harvesting time point and ran for the period of one cell duplication (i.e., simulation time = generation time), where one single cell was considered. Regarding the results for  $T_3$ , for example, the simulation began at time point  $T_3$  and the simulation time was 147 min. In biological sense, this means that at  $T_3$  one single cell starts to grow, divides after 147 min and ends



**TABLE 7 | Parameter for *in silico* analysis of ColE1-like plasmid replication control.**

Reaction kinetic	Constant	Low copy plasmid	High copy plasmid	Source
1	$k_1$ [ $M^{-1} \cdot \text{min}^{-1}$ ]	$1.7 \cdot 10^8$	Not available	a,-
2	$k_2$ [ $\text{min}^{-1}$ ]	0.17	Not available	a,-
3	$k_3$ [ $M^{-1} \cdot \text{min}^{-1}$ ]	$1.02 \cdot 10^8$	$3.05 \cdot 10^6$	c,c
4	$k_4$ [ $\text{min}^{-1}$ ]	48	20	a,c
5	$k_5$ [ $\text{min}^{-1}$ ]	12	12	a,a
6	$k_6$ [ $\text{min}^{-1}$ ]	4.3	4.3	a,a
7	$k_7$ [ $\text{min}^{-1}$ ]	3.8	4.19	f,f
8	$k_8$ [ $\text{min}^{-1}$ ]	4.3	4.3	a,a
9	$k_9$ [ $M^{-1} \cdot \text{min}^{-1}$ ]	0.25	0.25	a,a
10	$k_{10}$ [ $\text{min}^{-1}$ ]	44	44	a,a
11	$k_{11}$ [ $\text{min}^{-1}$ ]	0.085	0.085	a,a
12	$k_{12}$ [ $\text{min}^{-1}$ ]	17	17	a,a
13	$k_{13}$ [ $\text{min}^{-1}$ ]	34	Not available	a,-
14	$k_{14}$ [ $\text{min}^{-1}$ ]	$1.8 \cdot 10^5$	$1.8 \cdot 10^5$	b,b
15	$k_{15}$ [ $M^{-1} \cdot \text{min}^{-1}$ ]	$1.8 \cdot 10^5$	$1.8 \cdot 10^5$	b,b
16	$k_{16}$ [ $\text{min}^{-1}$ ]	12	12	g,g
17	$k_{17}$ [ $\text{min}^{-1}$ ]	0.25	0.25	g,g
18	$k_{18}$ [ $\text{min}^{-1}$ ]	6	6	a,a,d
19	$k_{19}$ [ $\text{min}^{-1}$ ]	0.25	0.25	a,a,d
20	$k_{20}$ [ $\text{min}^{-1}$ ]	4	Not available	a,-
21	$k_{21}$ [ $\text{min}^{-1}$ ]	0.35	0.35	a,a
22	$k_{22}$ [ $\text{min}^{-1}$ ]	0.35	0.35	a,a
23	$k_{23}$ [ $\text{min}^{-1}$ ]	0.14	Not available	a,-
24	$k_{24}$ [ $M \cdot \text{min}^{-1}$ ]	$7.99 \cdot 10^{-6}$	$7.99 \cdot 10^{-6}$	e,e
25	$k_{25}$ [ $M^{-1} \cdot \text{min}^{-1}$ ]	$1.02 \cdot 10^8$	$3.05 \cdot 10^6$	g,g
26	$k_{26}$ [ $\text{min}^{-1}$ ]	0.35	0.35	c,c

-No literature data is found for this value.

<sup>a</sup>Brendel and Perelson (1993).

<sup>b</sup>Wang et al. (2002) with the assumption that RNAI molecules are bound by tRNA molecules with the same kinetic constant like RNAI molecules bind to RNAI molecules in reaction 3.

<sup>c</sup>This work (parameter fitting).

<sup>d</sup> $k_{18}$  and  $k_{19}$  were calculated with an assumed RNA polymerase-transcription rate of  $50 \text{ nt s}^{-1}$  (von Hippel et al., 1984).

<sup>e</sup> $k_{24}$  was calculated with an assumed transcription rate of  $42 \text{ nt s}^{-1}$  [Gotta et al. (1991) and Klumpp (2011)].

<sup>f</sup> $k_7$  was calculated with an assumed DNA-polymerase-elongation rate of  $42 \text{ bp s}^{-1}$  (Alberts et al., 2007) under allowance of plasmid size.

<sup>g</sup>This work (estimated).

up with a PCN of 49, directly before the next cell division. The results of the parameter fitting and validation for *E. coli* DH5 $\alpha$ -pSUP 201-3 are depicted in **Table 8** in terms of simulated PCNs and experimentally determined PCNs. Besides the results assigned to time point  $T_2$ , which can be regarded as the outcome of the fitting procedure, the simulation results for the remaining time points  $T_1$  and  $T_3$  meet the experimental outcome very well. These results were achieved without any additional fitting steps, just with the parameters determined for time point  $T_2$ , which are summarized in **Table 7** (fitted parameters are assigned the letter “c”). Comparing the simulated PCNs and the measured PCNs of *E. coli* DH5 $\alpha$ -pSUP 201-3 and *E. coli* DH5 $\alpha$ -pCMV-lacZ (**Table 9**), respectively, one can conclude that the dynamic model supplied with this data set can sufficiently explain the bacterial behavior for different cultivation time points. Comparing the simulated PCNs and the measured PCNs of *E. coli* DH5 $\alpha$ -pSUP 201-3, it could be concluded that the dynamic model is able to reproduce the experimentally measured PCNs in an appropriate way. Looking at the simulated PCNs for *E. coli* DH5 $\alpha$ -pCMV-lacZ, it could be

**TABLE 8 | Results of the *in vitro* and *in silico* determined PCNs for DH5 $\alpha$ -pSUP.**

Time point	Measured pDNA [molecules/cell]	Simulated pDNA [molecules/cell]
$T_1$	$46 \pm 26$	49
$T_2$	$48 \pm 30$	47
$T_3$	$46 \pm 17$	49

The *in vitro* determined PCNs are the results of the plasmid concentration measurements done by the Plasmid Factory (Bielefeld, Germany) and the *in silico* determined PCNs are calculated by the dynamic model. As initial conditions for the simulations, the experimentally determined data for the three harvesting time points  $T_1$ ,  $T_2$ , and  $T_3$  were chosen. To these belong the measured RNAI-, RNAII-, and plasmid concentrations (**Table 5**) and the respective growth rates (**Table 3**). The simulation time was equal to the generation time (**Table 3**).

**TABLE 9 | Results of the *in vitro* and *in silico* determined PCNs for DH5 $\alpha$ -pCMV-lacZ.**

Time point	Measured pDNA [molecules/cell]	Simulated pDNA [molecules/cell]
$T_1$	$1514 \pm 1301$	1718
$T_2$	$2403 \pm 713$	2405
$T_3$	$5806 \pm 4828$	5421

The *in vitro* determined PCNs are the results of the plasmid concentration measurements done by the Plasmid Factory (Bielefeld, Germany) and the *in silico* determined PCNs are calculated by the dynamic model. As initial conditions for the simulations, the experimentally determined data for the three harvesting time points  $T_1$ ,  $T_2$ , and  $T_3$  were chosen. To these belong the measured RNAI-, RNAII-, and plasmid concentrations (**Table 6**) and the respective growth rates (**Table 4**). The simulation time was equal to the generation time (**Table 4**).

observed that in this case the model is also able to reproduce the experimentally determined PCNs in an appropriate way (**Table 9**). The simulated PCNs reside within the margin represented by the SD. To sum up, the simulation of the plasmid replication control for all harvesting time points is possible and the model is able to reproduce the PCN in the requested correctness.

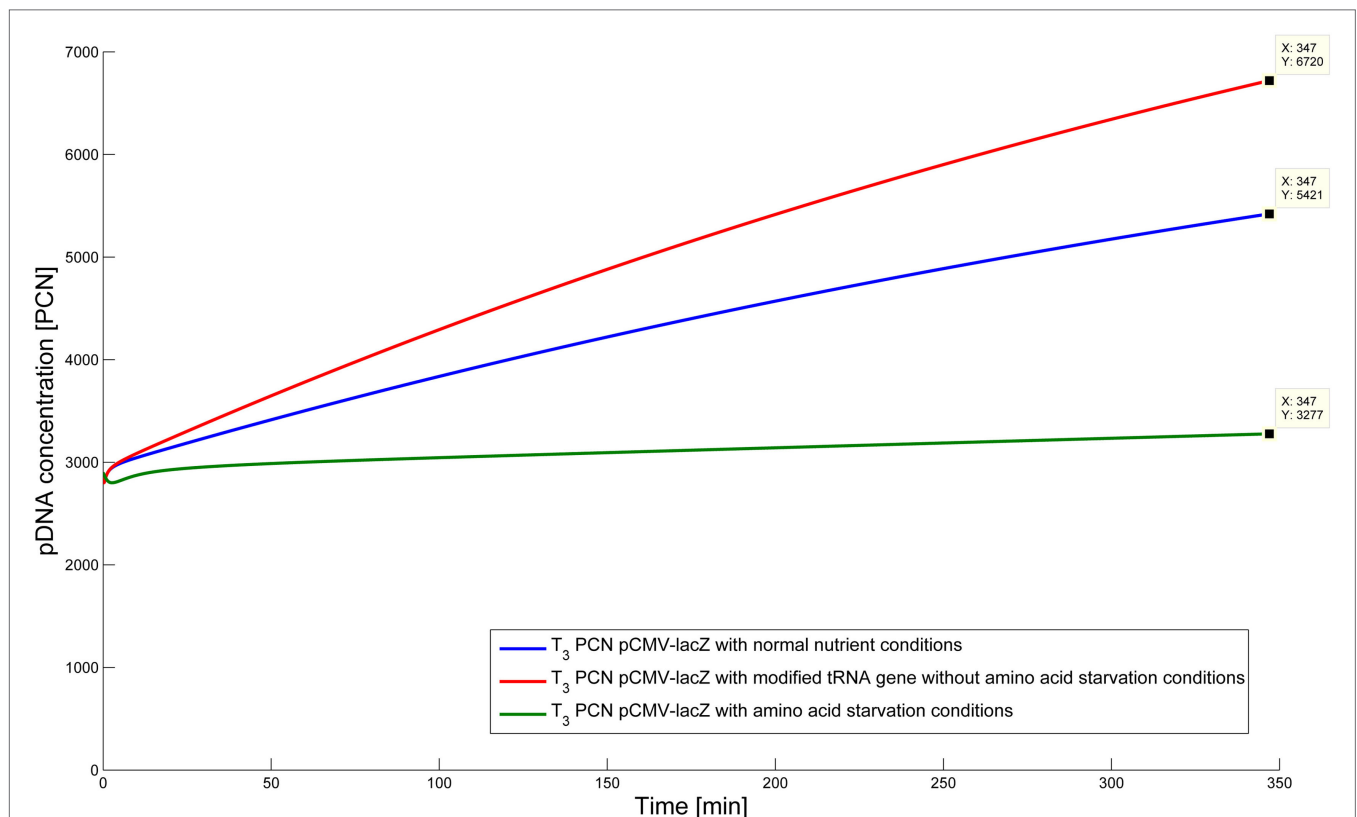
## Predictive Simulations

In this study, a dynamic model was developed to investigate the effect of uncharged tRNA molecules on the ColE1-like plasmid replication control. From this, the question arises if tRNA molecules can be modified in such a way that they cannot be charged any longer, and could we envisage an increase in plasmid production? In general, a high amount of uncharged tRNA molecules can be observed when the cell is afflicted with amino acid starvation. Under that condition, the protein synthesis is negatively affected and the majority of tRNA molecules in the cytoplasm are uncharged. Wróbel and Węgrzyn (1998) could show that a high amount of uncharged tRNA molecules is connected with an increase in the PCN (Wróbel and Węgrzyn, 1998). Thus, for plasmid production, starvation conditions would be advantageous. There is just one single problem: for a long-term production process, these conditions are not applicable, because the cells will not survive until the plasmid production cultivation is finished. One possible solution could be the overexpression of a tRNA gene. But in that case, there is a certain probability that the higher amount of uncharged tRNA will be rapidly charged by the amino-acyl- tRNA synthetases before they could influence the

plasmid replication. To overcome this problem, a modification of the tRNA molecule is required, which would preserve the charging and still positively influence the plasmid replication. In that case, the cell would be still vital under non-starvation conditions. But could such modified tRNA molecules actually push the plasmid replication? Therefore, the dynamic model was established, which enables to investigate how the plasmid replication control would behave, if a gene encoding a modified tRNA molecule was introduced into the genome.

To test the hypothesis that an uncharged tRNA would increase the PCN, the ColE1-like plasmid replication control of *E. coli* DH5 $\alpha$ -pCMV-lacZ was simulated for the following three conditions: growth under normal nutrient conditions, growth under amino acid starvation conditions, and growth under the influence of an inserted gene encoding for a modified tRNA under normal nutrient conditions. *E. coli* DH5 $\alpha$ -pCMV-lacZ was chosen, since for plasmid production usually high copy plasmids are used. The simulations could also be conducted for *E. coli* DH5 $\alpha$ -pSUP 201-3, leading qualitatively to the same result (data not shown). The simulations were done for all three conditions, beginning at  $T_3$ , for a period of one cell duplication [i.e., simulation time = generation time ( $T_3$ )]. One single cell is considered, which starts to grow at time point  $T_3$  and divides after 347 min. The initial conditions were the experimentally determined data at  $T_3$  (see **Tables 4** and **6**), which includes the RNAI- and RNAII

concentrations together with half of the respective determined plasmid concentrations and the appropriate measured growth rate. All unknown substrate concentrations were assigned the value 0. This means that the simulations consider one single cell, which starts to grow at time point  $T_3$  and divides after 347 min. In the moment of cell division, the simulation stops, so the daughter cells are not considered anymore. The time point  $T_3$  was chosen exemplarily, because at this time point the highest PCN was measured. Of course, it is possible to do these simulations just as well for  $T_1$  and  $T_2$ . The simulation results of the three mentioned conditions are graphically shown in **Figure 4**. The first simulation (**Figure 4**) considers growth conditions with normal nutrient supply. To describe this situation mathematically, the kinetic constant of uncharged tRNA synthesis ( $k_{24}$ ) was multiplied with the factor 0.01 to keep the amount of free uncharged tRNA molecules very low, because the amino acids supply is sufficient, so the amount of free uncharged tRNA molecules is very low. The factor 0.01 was chosen due to the assumption that under sufficient nutrient supply, only 1% of the tRNA molecules in the cell are uncharged. The *in silico* analysis showed that for a cell, growing under normal nutrient conditions, the PCN undergoes an increase to about 2523 (5421 plasmids at  $T_3$  – initial PCN = 2523). The second simulation (**Figure 4**) was done for a cell, which grows under amino acid starvation conditions. Here, the assumption was made, that as



**FIGURE 4 |** Simulation of the plasmid replication control for *E. coli* DH5 $\alpha$ -placZ, beginning at time point  $T_3$ , for the following three different growth conditions: blue, growth with normal

nutrient conditions; green, growth with amino acid starvation conditions, red, growth with modified tRNA gene without amino acid starvation conditions.

a consequence of insufficient nutrient supply, the intracellular protein concentrations are decreased and so all reactions run with only 10% reaction velocity (estimated). Mathematically spoken, every kinetic constant and the growth rate were multiplied by the factor 0.1. Furthermore, the factor of 0.01 for  $k_{24}$  was replaced by 1, because of the high amount of uncharged tRNA molecules under amino acid starvation. The second simulation predicts that within a period of cell duplication, one single cell produces only 379 plasmids. In the third simulation (**Figure 4**), the cells grow under normal nutrient supply with a high amount of modified tRNA molecules, which were encoded by an inserted modified tRNA gene. As for the mathematical description, this means that all reactions are running at 100%, because there is no amino acid starvation. Furthermore, the factor 0.01 for the kinetic constant  $k_{24}$  was again replaced by 1, since there are a lot of free modified tRNA molecules in the cell. Under these conditions, the model predicts a plasmid production of 3822 plasmids for one single cell within a period of cell duplication. Comparing all three simulations (**Figure 4**), it is apparent, that the plasmid production is lowest under insufficient nutrient supply. The highest plasmid production is predicted for the case when a gene, which encodes a modified, uncharged tRNA, is introduced into the genome.

## Discussion

In the last 10 years, the number of clinical trials in the field of human gene therapy has increased from 500 to 1500 (Prazeres, 2011). One can assume that once the clinical trials are completed and plasmid biopharmaceuticals enter the market, the demand of pDNA will increase. Since ColE1-type plasmids are often the basis for DNA vaccines or gene therapy products, it is important to understand how their replication is regulated (Prather et al., 2003). The dynamic model presented here is a valuable contribution to the modeling work done in this field, since it is more comprehensive than previous models (Ataai and Shuler, 1986; Bremer and Lin-Chao, 1986; Keasling and Palsson, 1989a,b; Brendel and Perelson, 1993; Wang et al., 2002). It extends the model proposed by Brendel and Perelson (1993), because it does not only incorporate the regulation through inhibitory RNAI molecules, but it considers also the control by uncharged tRNA molecules. Some previous models have focused on the PCN deviation from steady states derived from RNAI and RNAII concentrations (Paulsson et al., 1998), plasmid stability or probability of plasmid loss (Paulsson and Ehrenberg, 1998), as well as the formation of stable RNAII structures (Gulyaev et al., 1995). However, the mentioned reports did not show the effect of RNAI, RNAII, or uncharged tRNA on PCN linked to cell growth. Thus, the underlying reaction network of the proposed model is more detailed and the number of application conditions is increased. The simulation data are confirmed by *in vitro* measurements, obtained at three different time points. Since the data for each time point were measured for each time point in the same strain, they can be treated as homogeneous data sets. Many published dynamic models are based on heterogeneous data sets that consist of experimental data obtained from different strains and partly from different organisms (Klumpp, 2011). These heterogeneous

data sets are not optimal for modeling, because usually the model should help to investigate a specific organism, so data measured in a foreign organism might be unsuitable. In comparison to other growth rate measurements both strains used in this work grow very slowly. This slow growth could be reasoned by the usage of minimal medium for cultivation, because it contains only essential additives. On LB-medium both strains show the typical growth behavior for *E. coli* DH5 $\alpha$ .

The intracellular RNAI and RNAII concentrations were calculated for *E. coli* DH5 $\alpha$ -pSUP 201-3 and *E. coli* DH5 $\alpha$ -pCMV-lacZ via qRT-PCR with the aid of straight calibration lines. The intracellular RNAI- and RNAII concentrations measured for *E. coli* DH5 $\alpha$ -pSUP 201-3 decrease with proceeding growth of the bacterial culture. Furthermore, it was observed that there are more free RNAI molecules than RNAII molecules at each harvesting time point. The RNA measurements of this study are similar to the RNAI- and RNAII measurements done by Brenner and Tomizawa (1991) 23 years ago. Brenner and Tomizawa (1991) quantified the RNAI and RNAII concentrations densitometrically for an *E. coli* C600 derivative (SA791) strain carrying a pC<sub>er</sub> plasmid or a p $\Delta$ P4 plasmid (both low copy plasmids) applying quantitative probe protection experiments. They measured an average RNAII molecule number of  $1.9 \pm 0.9$  to  $3.7 \pm 1.0$  per cell and an average RNAI molecule number of  $333 \pm 49$  to  $499 \pm 36$  per cell (Brenner and Tomizawa, 1991). The RNAI concentrations determined by Brenner and Tomizawa (1991) are an order of 10 higher than the molecule numbers measured in this work. This could be reasoned by the different measuring methods. Possibly, the qRT-PCR is a more sensitive method compared to the quantification by quantitative probe protection experiments. Furthermore, Brenner and Tomizawa (1991) used a minimal medium supplemented with non-essential additives for cultivation, which could also have influenced the RNAI concentration. In comparison to *E. coli* DH5 $\alpha$ -pSUP 201-3, the RNAI- and RNAII concentrations measured for *E. coli* DH5 $\alpha$ -pCMV-lacZ do not decrease with proceeding growth of the bacterial culture. On the contrary, the intracellular RNAII concentrations increased with proceeding cultivation. Regarding the RNAI amount, the highest RNAI concentration was measured at time point  $T_2$  and the lowest at time point  $T_3$ . Since there are no comparative values published so far, the measured RNA concentrations of *E. coli* DH5 $\alpha$ -pCMV-lacZ cannot be compared with previously determined values. For *E. coli* DH5 $\alpha$ -pSUP 201-3, PCNs of 46 or 48 were measured at all three time points. The published PCNs for pBR322 plasmids are 15–20 plasmids per cell (Cooper and Cass, 2004), so the measured PCNs for *E. coli* DH5 $\alpha$ -pSUP 201-3 are around two times higher. But considering the SDs of the measurements of this study, it could be seen that the PCNs measured by Cooper and Cass (2004) reside within the same range, defined by the SDs. Again the difference could be explained by the cultivation in minimal medium with no non-essential supplements used in this work, because the *E. coli* DH5 $\alpha$ -pSUP 201-3-cells grew very slowly and a slow growth promotes the intracellular plasmid concentration (Bremer and Lin-Chao, 1986). Taken as a whole, the plasmid amount per cell of *E. coli* DH5 $\alpha$ -pSUP 201-3 is stable within all three time points. In contrast to the low copy plasmid strain *E. coli* DH5 $\alpha$ -pSUP

201-3, the plasmid concentration of *E. coli* DH5 $\alpha$ -pCMV-lacZ increases with progressing growth of the bacterial culture. The highest plasmid concentration was measured at time point  $T_3$  with  $5806 \pm 4828$  plasmids per cell. It can be assumed that the increasing RNAPII concentration is responsible for the increase of the PCN, because more RNAPII molecules are available to serve as a primer. Comparing these measured PCNs with literature values shows that they are very high. For pUC plasmids, copy numbers of 500–700 were reported (Cooper and Cass, 2004). To address these differences, plasmid numbers were confirmed by qRT-PCR, resulting in the same order of magnitude. As for *E. coli* DH5 $\alpha$ -pSUP 201-3, the minimal medium and the slow growth could be responsible for that high plasmid concentrations.

After the experimental data had been introduced to the dynamic model, the unknown parameters were partly obtained by parameter fitting. After comparing simulation results with our own experimental as well as literature data and achieving a satisfying validation, the model was ready to be used to do *in silico* analysis. The hypothesis that modified tRNA molecules would have a similar pushing effect on plasmid production like uncharged tRNA molecules have under starvation conditions was tested. The simulations showed that under amino acid starvation, the smallest amount of plasmids is produced. As for the effect of modified tRNA molecules on plasmid production, the model predicts an increase compared to the plasmid production under normal nutrient conditions. Thus, the insertion of a gene encoding modified uncharged tRNA molecules would have a positive effect on the plasmid production. Since the hypothesis is confirmed by *in silico* analysis with the established model, the next step will be to concentrate on the design of a modified uncharged tRNA gene. This gene should encode for a modified tRNA molecule, which will not be recognized by an amino-acyl synthetase and therefore will not be charged anymore. Additionally, all elements influencing the ColE1-like plasmid replication control have to maintain functionality.

## References

- Alberts, B., Johnson, A., Walter, P., and Lewis, J. (2007). *Molecular Biology of the Cell*, 5th Revised Edn. New York & Abington: Garland Science, Taylor & Francis.
- Ataai, M. M., and Shuler, M. L. (1986). Mathematical model for the control of ColE1 type plasmid replication. *Plasmid* 16, 204–212. doi:10.1016/0147-619X(86)90058-2
- Bower, D. M., and Prather, K. L. J. (2009). Engineering of bacterial strains and vectors for the production of plasmid DNA. *Appl. Microbiol. Biotechnol.* 82, 805–813. doi:10.1007/s00253-009-1889-8
- Bremer, H., and Lin-Chao, S. (1986). Analysis of the physiological control of replication of ColE1-type plasmids. *J. Theor. Biol.* 123, 453–470. doi:10.1016/S0022-5193(86)80213-2
- Brendel, V., and Perelson, A. S. (1993). Quantitative model of ColE1 plasmid copy number control. *J. Mol. Biol.* 229, 860–872. doi:10.1006/jmbi.1993.1092
- Brenner, M., and Tomizawa, J.-I. (1991). Quantitation of ColE1-encoded replication elements. *Proc. Natl. Acad. Sci. U.S.A.* 88, 405–409. doi:10.1073/pnas.88.2.405
- Cooper, J., and Cass, T. (2004). *Biosensors – Practical Approach*, 2nd Edn. Oxford, UK: Oxford University Press.
- del Solar, G., and Espinosa, M. (2000). Plasmid copy number control: an ever-growing story. *Mol. Microbiol.* 37, 492–500. doi:10.1046/j.1365-2958.2000.02005.x
- Funahashi, A., Matsuoka, Y., Jouraku, A., Morohashi, M., Kikuchi, N., and Kitano, H. (2008). CellDesigner 3.5: a versatile modeling tool for biochemical networks. *Proc. IEEE* 96, 1254–1265. doi:10.1109/JPROC.2008.925458
- Gotta, S. L., Miller, O. L., and French, S. L. (1991). rRNA transcription rate in *Escherichia coli*. *J. Bacteriol.* 173, 6647–6649.

## Conclusion

In this study, a dynamic model of ColE1-like plasmid replication control for a low copy plasmid and a high copy plasmid was established. The model comprises the replication control by regulatory RNA molecules and by uncharged tRNA molecules. The big advantage of this model is that it is confirmed by *in vitro* measurements, which were done for the same strains and at three different time points. The established model was used to simulate the plasmid replication control for the period of one cell duplication, where one single cell was considered. Nevertheless, the model can also be used to predict the plasmid replication control for more than one period of cell duplication, which could be conducted for constant growth conditions as well as for varying growth conditions. The simulations done with the dynamic model within this work predict that inserting a gene encoding modified uncharged tRNA molecules would increase the PCN. Since the overall aim is to improve the plasmid production to accommodate the prospective demand on pDNA, this predicted effect of modified tRNA molecules is very important. For that reason, further investigation should be concentrated on the construction of such a modified uncharged tRNA gene.

## Acknowledgments

The Graduate Cluster Industrial Biotechnology (CLIB2021) and the BMBF (01DN12101) are highly acknowledged for financially supporting this project. We acknowledge support for the Article Processing Charge by the Deutsche Forschungsgemeinschaft and the Open Access Publication Fund of Bielefeld University.

## Supplementary Material

The Supplementary Material for this article can be found online at <http://journal.frontiersin.org/article/10.3389/fbioe.2015.00127>

- Grabherr, R., and Bayer, K. (2002). Impact of targeted vector design on ColE1 plasmid replication. *Trends Biotechnol.* 20, 257–260. doi:10.1016/S0167-7799(02)01950-9
- Gulyaev, A. P., van Batenburg, F. H., and Pleij, C. W. (1995). The influence of a metastable structure in plasmid primer RNA on antisense RNA binding kinetics. *Nucleic Acids Res.* 23, 3718–3725. doi:10.1093/nar/23.18.3718
- Itoh, T., and Tomizawa, J.-I. (1980). Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. *Proc. Natl. Acad. Sci. U.S.A.* 77, 2450–2454. doi:10.1073/pnas.77.5.2450
- Keasling, J. D., and Palsson, B. O. (1989a). On the kinetics of plasmid replication. *J. Theor. Biol.* 136, 487–492. doi:10.1016/S0022-5193(89)80161-4
- Keasling, J. D., and Palsson, B. O. (1989b). ColE1 plasmid replication: a simple kinetic description from a structured model. *J. Theor. Biol.* 141, 447–461. doi:10.1016/S0022-5193(89)80230-9
- Klumpp, S. (2011). Growth-rate dependence reveals design principles of plasmid copy number control. *PLoS ONE* 6:e20403. doi:10.1371/journal.pone.0020403
- Kutzler, M. A., and Weiner, D. B. (2008). DNA vaccines: ready for prime time? *Nat. Rev. Genet.* 9, 776–788. doi:10.1038/nrg2432
- Lara, A. R., Caspeta, L., Gosset, G., Bolívar, F., and Ramírez, O. T. (2008). Utility of an *Escherichia coli* strain engineered in the substrate uptake system for improved culture performance at high glucose and cell concentrations: an alternative to fed-batch cultures. *Biotechnol. Bioeng.* 99, 893–901. doi:10.1002/bit.21664
- Lara, A. R., and Ramírez, O. T. (2012). Plasmid DNA production for therapeutic applications. *Methods Mol. Biol.* 824, 271–303. doi:10.1007/978-1-61779-433-9\_14



- Lin-Chao, S., Chen, W. T., and Wong, T. T. (1992). High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. *Mol. Microbiol.* 6, 3385–3393. doi:10.1111/j.1365-2958.1992.tb02206.x
- Mairhofer, J., and Lara, A. R. (2014). Advances in host and vector development for the production of plasmid DNA vaccines. *Methods Mol. Biol.* 1139, 505–541. doi:10.1007/978-1-4939-0345-0\_38
- Paulsson, J., and Ehrenberg, M. (1998). Trade-off between segregational stability and metabolic burden: a mathematical model of plasmid ColE1 replication control. *J. Mol. Biol.* 279, 73–88. doi:10.1006/jmbi.1998.1751
- Paulsson, J., Nordström, K., and Ehrenberg, M. (1998). Requirements for rapid plasmid ColE1 copy number adjustments: a mathematical model of inhibition modes and RNA turnover rates. *Plasmid* 39, 215–234. doi:10.1006/plas.1998.1338
- Prather, K. J., Sagar, S., Murphy, J., and Chartrain, M. (2003). Industrial scale production of plasmid DNA for vaccine and gene therapy: plasmid design, production, and purification. *Enzyme Microb. Technol.* 33, 865–883. doi:10.1016/S0141-0229(03)00205-9
- Prazeres, D. M. F. (2011). *Plasmid Biopharmaceuticals – Basics, Applications and Manufacturing*, 1st Edn. New Jersey: Wiley Publishing Company.
- Schleef, M., and Blaes, M. (2009). “Chapter 25: production of plasmid DNA as a pharmaceutical,” in *Methods in Molecular Biology, Gene Therapy and Cancer*, Vol. 54, eds Walther W. and Stein U. S. (New York: Humana Press), 471–495.
- Schumann, W. (2001). “Chapter 1: the biology of plasmids,” in *Plasmids for Therapy and Vaccination*, ed. Schleef M. (Weinheim: Wiley-VCH), 1–28.
- Simon, R. S., Priefer, U., and Pühler, A. (1983). A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Nat. Biotechnol.* 1, 784–791. doi:10.1038/nbt1183-784
- Tomizawa, J.-I. (1984). Control of ColE1 plasmid replication: the process of binding of RNA I to the primer transcript. *Cell* 38, 861–870. doi:10.1016/0092-8674(84)90281-2
- Tomizawa, J.-I. (1985). Control of ColE1 plasmid replication: initial interaction of RNA I and the primer transcript is reversible. *Cell* 40, 527–535. doi:10.1016/0092-8674(85)90201-6
- Tomizawa, J.-I. (1986). Control of ColE1 plasmid replication: binding of RNA I to RNA II and inhibition of primer formation. *Cell* 47, 89–97. doi:10.1016/0092-8674(86)90369-7
- Tomizawa, J.-I. (1990). Control of ColE1 plasmid replication interaction of rom protein with an unstable complex formed by RNA I and RNA II. *J. Mol. Biol.* 212, 695–708. doi:10.1016/0022-2836(90)90231-A
- Tomizawa, J.-I., and Itoh, T. (1981). Plasmid ColE1 incompatibility determined by interaction of RNA I with primer transcript. *Proc. Natl. Acad. Sci. U.S.A.* 78, 6096–6100. doi:10.1073/pnas.78.10.6096
- von Hippel, P. H., Bear, D. G., Morgan, W. D., and McSwiggen, J. A. (1984). Protein-nucleic acid interactions in transcription: a molecular analysis. *Annu. Rev. Biochem.* 53, 389–446. doi:10.1146/annurev.bi.53.070184.002133
- Wang, Z., Le, G., Shi, Y., Węgrzyn, G., and Wróbel, B. (2002). A model for regulation of ColE1-like plasmid replication by uncharged tRNAs in amino acid-starved *Escherichia coli* cells. *Plasmid* 47, 69–78. doi:10.1006/plas.2001.1562
- Wróbel, B., and Węgrzyn, G. (1998). Replication regulation of ColE1-like plasmids in amino acid-starved *Escherichia coli*. *Plasmid* 39, 48–62. doi:10.1006/plas.1997.1319
- Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains; nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33, 103–119. doi:10.1016/0378-1119(85)90120-9
- Yavachev, L., and Ivanov, I. (1988). What does the homology between *E. coli* tRNAs and RNAs controlling ColE1 plasmid replication mean? *J. Theor. Biol.* 131, 235–241. doi:10.1016/S0022-5193(88)80240-6

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Freudenau, Lutter, Baier, Schleef, Bednarz, Lara and Niehaus. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Shikimic acid production in *Escherichia coli*: from classical metabolic engineering strategies to omics applied to improve its production

## OPEN ACCESS

### Edited by:

Hilal Taymaz Nikerel,  
Bogazici University, Turkey

### Reviewed by:

Maria Suarez Diez,  
Helmholtz Zentrum für  
Infektionsforschung, Germany  
Hannes Link,  
Max Planck Institute for Terrestrial  
Microbiology, Germany

### \*Correspondence:

Adelfo Escalante,  
Departamento de Ingeniería Celular y  
Biocatálisis, Instituto de  
Biotecnología, Universidad Nacional  
Autónoma de México,  
Avenida Universidad 2001,  
Colona Chamilpa, Cuernavaca,  
Morelos 62210, Mexico  
adelfo@ibt.unam.mx

### Specialty section:

This article was submitted to  
Systems Biology, a section of the  
journal *Frontiers in Bioengineering  
and Biotechnology*

**Received:** 08 July 2015

**Accepted:** 07 September 2015

**Published:** 23 September 2015

### Citation:

Martínez JA, Bolívar F and  
Escalante A (2015) Shikimic acid  
production in *Escherichia coli*: from  
classical metabolic engineering  
strategies to omics applied to  
improve its production.  
*Front. Bioeng. Biotechnol.* 3:145.  
doi: 10.3389/fbioe.2015.00145

**Juan Andrés Martínez, Francisco Bolívar and Adelfo Escalante\***

Departamento de Ingeniería Celular y Biocatálisis, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Mexico

Shikimic acid (SA) is an intermediate of the SA pathway that is present in bacteria and plants. SA has gained great interest because it is a precursor in the synthesis of the drug oseltamivir phosphate (OSF), an efficient inhibitor of the neuraminidase enzyme of diverse seasonal influenza viruses, the avian influenza virus H5N1, and the human influenza virus H1N1. For the purposes of OSF production, SA is extracted from the pods of Chinese star anise plants (*Illicium* spp.), yielding up to 17% of SA (dry basis content). The high demand for OSF necessary to manage a major influenza outbreak is not adequately met by industrial production using SA from plants sources. As the SA pathway is present in the model bacteria *Escherichia coli*, several “intuitive” metabolically engineered strains have been applied for its successful overproduction by biotechnological processes, resulting in strains producing up to 71 g/L of SA, with high conversion yields of up to 0.42 (mol SA/mol Glc), in both batch and fed-batch cultures using complex fermentation broths, including glucose as a carbon source and yeast extract. Global transcriptomic analyses have been performed in SA-producing strains, resulting in the identification of possible key target genes for the design of a rational strain improvement strategy. Because possible target genes are involved in the transport, catabolism, and interconversion of different carbon sources and metabolic intermediates outside the central carbon metabolism and SA pathways, as genes involved in diverse cellular stress responses, the development of rational cellular strain improvement strategies based on omics data constitutes a challenging task to improve SA production in currently over-producing engineered strains. In this review, we discuss the main metabolic engineering strategies that have been applied for the development of efficient SA-producing strains, as the perspective of omics analysis has focused on further strain improvement for the production of this valuable aromatic intermediate.

**Keywords:** *Escherichia coli*, metabolic engineering, shikimic acid, transcriptome, metabolome, antiviral drug, influenza

## Introduction

Compounds derived from the aromatic amino acid (AA) pathway play important roles in the pharmaceutical and food industries as raw materials, additives, or final products (Patnaik et al., 1995; Bongaerts, 2001; Báez et al., 2001; Yi et al., 2002; Chandran et al., 2003; Báez-Viveros et al., 2004; Gosset, 2009). This metabolic pathway is present in bacteria and plants, starting with condensation of the central carbon metabolism (CCM) intermediates phosphoenolpyruvate (PEP) and erythrose-4-phosphate (E4P) to form the first AA pathway intermediate D-arabinoheptulosonate-7-phosphate (DAHP). From this compound to chorismic acid (CHA), the pathway is mostly linear and represents the first part of the AA pathway, known as the common AA pathway or the shikimic acid (SA) pathway (**Figure 1**). One of the specific intermediates on this pathway is SA, which is a highly functionalized six-carbon cyclic compound with three asymmetric centers. Therefore, SA is an enantiomeric precursor for the production of many high valuable biological active compounds for different industries. SA is the precursor for the synthesis of compounds with diverse pharmaceutical applications, including as an antipyretic, antioxidant, anticoagulant, antithrombotic, anti-inflammatory, or analgesic agent, for the synthesis of anticancer drugs, such as (+)-zeylenone (which has been shown to inhibit nucleoside transport in Ehrlich carcinoma cells and to be cytotoxic to cultured cancer cells), and for antibacterial or hormonal applications [reviewed in Estevez and Estevez (2012), Liu et al. (2012), and Diaz Quiroz et al. (2014)] (**Figure 2**).

Specifically, SA has great pharmaceutical relevance because it is the precursor for the chemical synthesis of oseltamivir phosphate (OSF), known as Tamiflu®, used as the antiviral inhibitor of the neuraminidase enzyme for the treatment of diverse seasonal influenza viruses, including influenza A and B, the avian influenza virus H5N1, and the human influenza virus H1N1 (Krämer et al., 2003; Estevez and Estevez, 2012; Ghosh et al., 2012; Diaz Quiroz et al., 2014). For this purpose, SA is obtained from the seed of the Chinese star anise plant *Illicium verum*, which contains between 2 and 7% of the intermediate. However, it can only be retrieved from plants after 6 years of crop growth and harvested in September and October (Li et al., 2007; Raghavendra et al., 2009; Wang et al., 2011). To recover SA from the seed, a 10-step process is required, taking ~30 kg of seed to produce 1 kg of SA. According to Li et al. (2007) on their 2007 patent, ~90% of the Chinese harvest is used by Roche (2009) for OSF production.

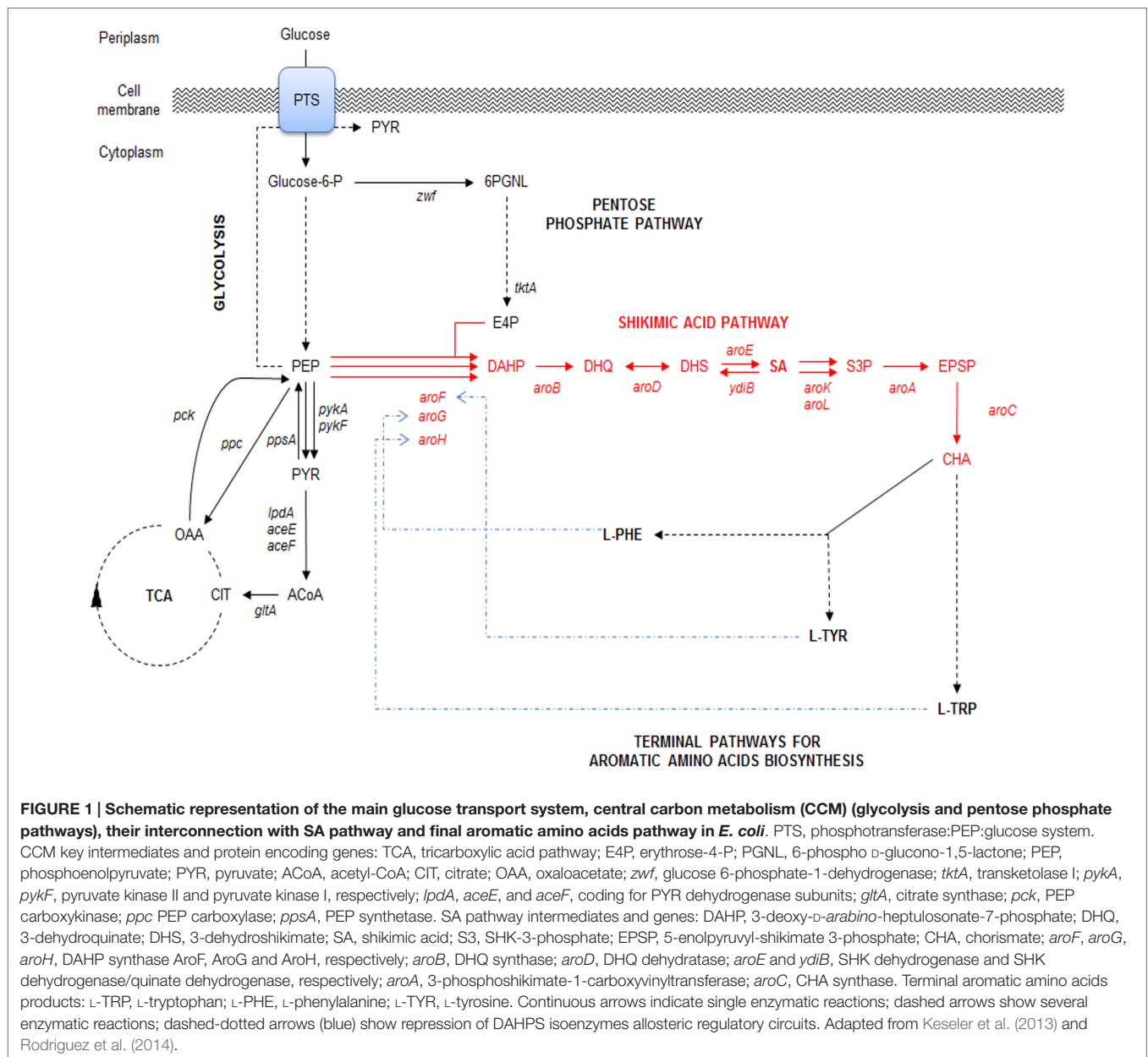
In 2009, Roche reported Tamiflu® sales to be 3.5 billion dollars, with a production capacity of up to 33 million treatments per month and 400 million packages per year (Scheiwiller and Hirschi, 2010). For the antiviral production, up to 1.3 g of SA are required to manufacture 10 doses to treat only one person, estimating a production requirement for this antiviral drug alone of ~520,000 kg/year (Rangachari et al., 2013). Even so, this reported production capacity could be insufficient in the case of an influenza pandemic, particularly with more pathogenic and infective strains. An estimated production of 30 billion doses, requiring 3.9 million kilograms of SA, would be necessary to cover a severe influenza outbreak (Rangachari et al., 2013).

According to the World Health Organization regarding influenza outbreak preparedness, only 66 million people in medium to low income countries are covered up with antiviral stocks, representing only 2.25% of the populations in these countries (World Health Organization, 2011). This situation results in a possibly low production capacity since in 2010, 100 million people were infected with common strains of influenza in Europe, Japan, and the United States alone. Moreover, before 2010, pandemic influenza has affected between 20 and 40% of the population, causing over 20 million deaths (Scheiwiller and Hirschi, 2010; World Health Organization, 2011).

For the reasons mentioned before and due to the relevance of SA in diverse industrial setups, many studies concerning SA production have been conducted within the past years, resulting in new and insightful strategies for its production, including recovery technologies, chemical synthesis methods, and biotechnological production methods using microorganisms. In fact, one of the most studied alternatives for SA production processes is biotechnological synthesis using recombinant microbial strains that are capable of producing high yields and that have high productivities, as there are key advantages over chemical synthesis, which include environmental friendliness, the availability and abundance of low-cost renewable feed stocks, and selectivity and diversity of the obtained products (Chen et al., 2013). These strains can be obtained by genetic modification, altering cellular properties to enhance their production capacity through the application of diverse metabolic engineering (ME) approaches (Krämer et al., 2003; Ghosh et al., 2012; Diaz Quiroz et al., 2014). However, despite the great achievements accomplished through this discipline, performance improvement has become limited after the first breakthroughs, mainly because of the traditional local pathway modification strategies. This is probably due to the limited understanding of the overall mechanism of metabolic regulation (Matsuoka and Shimizu, 2012). Therefore, given the importance of finding not only a particular pathway but also global information regarding cell physiology and metabolism to overcome production limitations, a systems biology approach supported by omics data may be the solution for improving SA production. The goal of this work is not only to review the literature on the great biotechnological achievements made for SA production, mainly in *Escherichia coli*, but also to outline future perspectives on research performed in the omics era, which could provide relevant tools for understanding cell behavior and production optimization via biotechnological processes.

## Classical Metabolic Engineering Approaches for SA Production

Metabolic engineering has been used since 1991 for strain modification by using recombinant DNA technology to enhance the production of specific metabolites (Matsuoka and Shimizu, 2012). The efforts to use ME have extended from the early years to optimize many cellular behaviors or parameters, such as substrate consumption, robustness, and tolerance toward toxic compounds and media conditions (Matsuoka and Shimizu, 2012). Classical ME strategies for strain development include various steps, such as the selection of a proper organism, elimination of competing pathways, deregulation of desired pathways at the enzyme activity

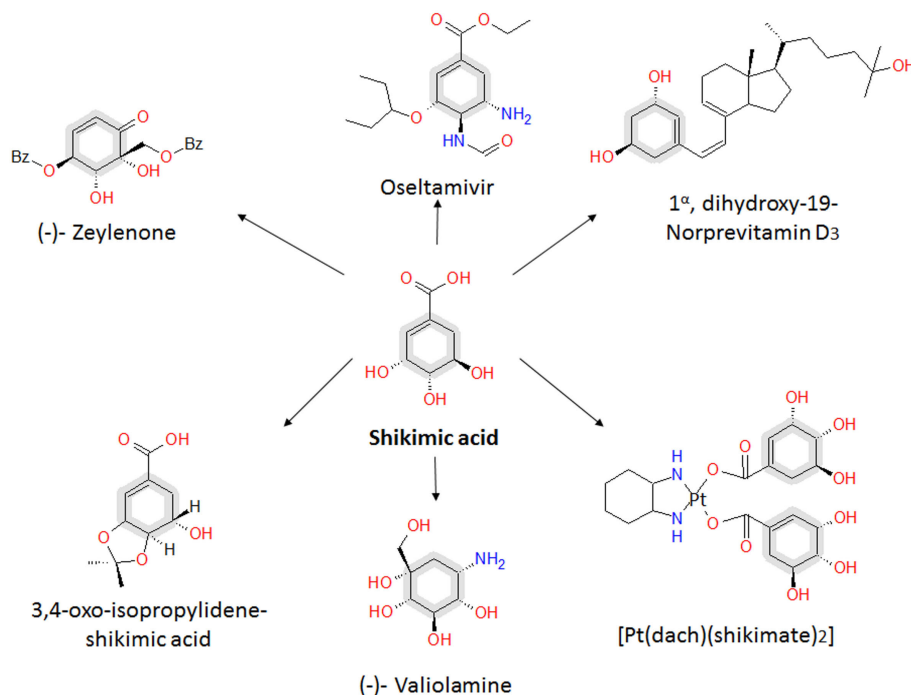


and transcriptional levels, and overexpression of enzymes at flux bottlenecks (Patnaik et al., 1995). Regarding the selection of an organism, *E. coli* has been preferred for industrial purposes and ME applications because of the knowledge available on *E. coli* physiology and the great numbers of tools developed to modify its genome (Chen et al., 2013). Therefore, many advances had been made regarding SA in *E. coli*, rendering strains capable of being used in industrial applications (Frost et al., 2002; Li et al., 2007).

In *E. coli*, the SA pathway starts by condensation of the CCM intermediates PEP and E4P by three DAHP synthase isoenzymes, AroG, AroF, and AroH (coded by *aroG*, *aroF*, and *aroH*, respectively), to produce DAHP. These three isoenzymes are responsible for the redirection from CCM intermediates toward the synthesis of aromatic compounds and are allosterically regulated specifically by the final products of AA biosynthesis. AroG catalyzes

~80% of DAHPS activity and is specifically feedback regulated by L-phenylalanine, AroF (~20% of DAHPS activity) is feedback regulated by L-tyrosine, and AroH (~1% of DAHPS activity) is regulated by L-tryptophan. Additionally, the transcription of *aroG* and *aroF* is controlled by the *tyrR* repressor, with the end products of the AA pathway (L-phenylalanine and L-tyrosine, respectively) acting as corepressors, whereas the transcription of *aroH* is controlled by the *trpP* repressor, with L-tryptophan acting as a corepressor (Keseler et al., 2013) (Figure 1). The ME solution for this first flux bottleneck is the expression of a DAHP AroG and AroF synthase that is not sensitive to feedback inhibition (*fbr*) (AroG<sup>fbr</sup> and AroF<sup>fbr</sup>). Mutations in the *aroG* and *aroF* genes lead to L-phenylalanine and L-tyrosine feedback-insensitive mutants with increased net carbon flux from CCM to the SA pathway (Keseler et al., 2013; Lin et al., 2014; Rodriguez et al., 2014); these





**FIGURE 2 | Relevant SA derivatives with high added value.** OSF, viral inhibitor of diverse influenza virus types, including seasonal types A and B, avian virus H5N1, and human virus H1N1. (–)-Zeylenone is a compound with antiviral, anticancer, and antibiotic activities. (–)-Valiolamine, a very strong  $\alpha$ -glucosidase with inhibitory activity against porcine intestinal enzymes sucrase, maltase, and isomaltase. [Pt(dach)(SA)<sub>2</sub>] is an active compound against L1210 leukemia. 3,4-Oxo-isopropylidene-SA, with antithrombotic activity and anti-inflammatory effects. Analogs of 1 $\alpha$ , dihydroxy-19-Nor previtamin D3 is a compound with promising applications in the treatment of osteoporosis and malignancies. Adapted from Estevez and Estevez (2012) and Diaz Quiroz et al. (2014).

mutants have been used in most SA production strains (Table 1) (Chandran et al., 2003; Escalante et al., 2010; Chen et al., 2012, 2014; Rodriguez et al., 2013). Forward reactions convert DAHP to dehydroquinic acid (DHQ), then to 3-dehydroquinate (DHS) and finally to SA by the enzymes 3-dehydroquinate synthase (*aroB*), 3-dehydroquinate dehydratase (*aroD*), and shikimate dehydrogenase (*aroE*), respectively (Figure 1). Although the pathway to SA conversion is small and linear, its regulation and the competition for precursor metabolites remain quite complicated because the SA pathway is dependent on the glycolytic and pentose phosphate pathways (PPPs) to provide the starting precursors PEP and E4P, respectively (Gosset, 2009; Escalante et al., 2012; Ghosh et al., 2012; Rodriguez et al., 2014).

SA production starts in CCM, further away from glucose consumption. In *E. coli*, the majority of glucose transport occurs via the PEP:glucose phosphotransferase system (PTS), which uses a phosphate group from one molecule of PEP to simultaneously import and phosphorylate periplasmic glucose, resulting into 6-phosphate glucose (G6P) and pyruvate (PYR) (Figure 1). For these reasons, the application of ME strategies only on the SA pathway would not render a significantly optimized strain for SA production (Ghosh et al., 2012).

To optimize productivity and yields from a given carbon source, modification of the CCM pathways supplying the needed precursors and energy sources for product synthesis is required (Patnaik and Liao, 1994). For the E4P supply, the PPP is the responsible for its production. The overexpression of transketolase I (TktA,

coded by *tktA*) and transaldolase (coded by *talA*), resulting in the preferential use of TktA to improve the E4P pool for the synthesis of DAHP (Flores et al., 1996; Draths et al., 1999; Frost et al., 2002; Chandran et al., 2003; Escalante et al., 2010; Rodriguez et al., 2013).

Regarding increasing the PEP pool, the first problem arises with the consumption of 50% of the PEP resulting from the catabolism of one molecule of glucose-6-P by PTS during the translocation and phosphorylation of one molecule of glucose. A rational approach is to reconvert PYR to PEP by overexpressing PEP synthase (coded by *pps*); this solution, along with the expression of a DAHPS<sup>fb</sup> (AroG<sup>fb</sup> or AroF<sup>fb</sup>), leads to a 51% (mol/mol) yield of DHS and related SA pathway metabolites. This yield is in fact higher than the 43% (mol/mol) yield calculated from stoichiometric reactions, reflecting the effective redistribution of the PEP to PYR pool ratio and the ability of the strain to redirect this new imbalance into the SA pathway (Yi et al., 2002; Chandran et al., 2003; Krämer et al., 2003; Escalante et al., 2010; Rodriguez et al., 2013). Overexpression of the *pps* gene has been studied; the maximum yield of SA is not obtained under the maximum concentration of the enzyme. In fact, it has been found that expression of this enzyme over the optimized level would only reduce the yields of SA intermediates, probably due to energetic imbalances (Yi et al., 2002).

The maximum theoretical yield limitation can be changed by restructuring the metabolic network, providing the system with a new stoichiometric matrix. Therefore, a natural solution for the

**TABLE 1 | Relevant *E. coli* engineered strains for SA production.**

Producing strain	Phenotypic traits	Special comments	Culture conditions	Title (g/L)	Yield (mol SA/mol glc)	Reference
SA112	BW25113 $\Delta$ aroKL, P <sub>pps</sub> ::P <sub>lac</sub> Q1, P <sub>csrB</sub> ::P <sub>lac</sub> Q1Pt5-pps, P <sub>TS</sub> -csrB, 5P <sub>tac</sub> -tktA	ClChE evolved to optimize SA production	Shake flasks cultures with 10 g/L glc, 1 g/L peptone	1.70	0.25	Cui et al. (2014)
DHPYAAS-T7	DH5 $\alpha$ $\Delta$ ptsH $\Delta$ lrr, $\Delta$ aroKL, $\Delta$ ydiB pAOC-TGEFB:aroE, aroB, glk, tktA, aroF <sup>thr</sup>	Plasmid overexpression of SA related genes	Shake flasks cultures (50 mL), M9 broth supplemented with 25 g/L glycerol, 10 g/L peptone, 15 g/L YE	1.066	0.23	Chen et al. (2012)
PB12.SA22	JM101 $\Delta$ ptsH, ptsI, crr::Km <sup>r</sup> $\Delta$ aroKL::cm <sup>r</sup> pJLBaroG <sup>thr</sup> tktA, pTOPO-aroBaroE	Laboratory evolved PTS <sup>-</sup> glucose <sup>-</sup> into glucose <sup>+</sup> derivative phenotype	1 L batch bioreactor 25 g/L glc and 15 g/L YE	7.05	0.22	Escalante et al. (2010)
SA5	B0013 $\Delta$ aroKL::dif $\Delta$ ptsG::dif $\Delta$ ydiB::dif $\Delta$ ackA-pta::dif pTH-aroG <sup>thr</sup> -ppsA-tktA	Plasmid over expression of SA related genes	7 L fed-batch bioreactor, initial 15 g/L glc supplemented with AA and vitamins	14.6	0.3	Chen et al. (2014)
SA114	BW25113 $\Delta$ aroKL, P <sub>pps</sub> ::P <sub>lac</sub> Q1, P <sub>csrB</sub> ::P <sub>lac</sub> Q1 P <sub>TS</sub> -pps, P <sub>TS</sub> -csrB, 5P <sub>tac</sub> -tktA, 5P <sub>tac</sub> -pntAB	ClChE evolved to optimize SA	Shake flasks, 10 g/L glc, 1 g/L peptone	2.99	0.31	Cui et al. (2014)
SP1.1pts-/pSC6.090B	DH5 $\alpha$ $\Delta$ ptsH-ptsI-crr $\Delta$ serA::aroB $\Delta$ aroL::Tn10 $\Delta$ aroK::Cm <sup>r</sup> P <sub>tac</sub> Glf glk, aroF <sup>thr</sup> tktA, P <sub>tac</sub> aroE, serA	Heterologous glk and glf from <i>Z. mobilis</i> to restore glc transport and phosphorylation in PTS <sup>-</sup> glc <sup>-</sup> phenotype	10 L fed-batch bioreactor 55–170 mM Glc + 15 g/L YE	84	0.33	Chandran et al. (2003)
SA116	BW25113 $\Delta$ aroKL, P <sub>pps</sub> ::P <sub>lac</sub> Q1, P <sub>csrB</sub> ::P <sub>lac</sub> Q1 P <sub>TS</sub> -pps, P <sub>TS</sub> -csrB, 5P <sub>tac</sub> -tktA, 5P <sub>tac</sub> -nadK	ClChE evolved to optimize SA	Shake flasks 10 g/L glc, 1 g/L peptone	3.12	0.33	Cui et al. (2014)
AR36	JM101 $\Delta$ ptsH, ptsI, crr::Km <sup>r</sup> $\Delta$ aroKL::cm <sup>r</sup> $\Delta$ pykF $\Delta$ lacI pTcAro6-aroB, tktA, aroG <sup>thr</sup> , aroE, aroD zwf	Constitutive strong over expression by synthetic operon on plasmid	1 L batch bioreactor 100 g/L glc + 15 g/L YE	41.8	0.42	Rodríguez et al. (2013)

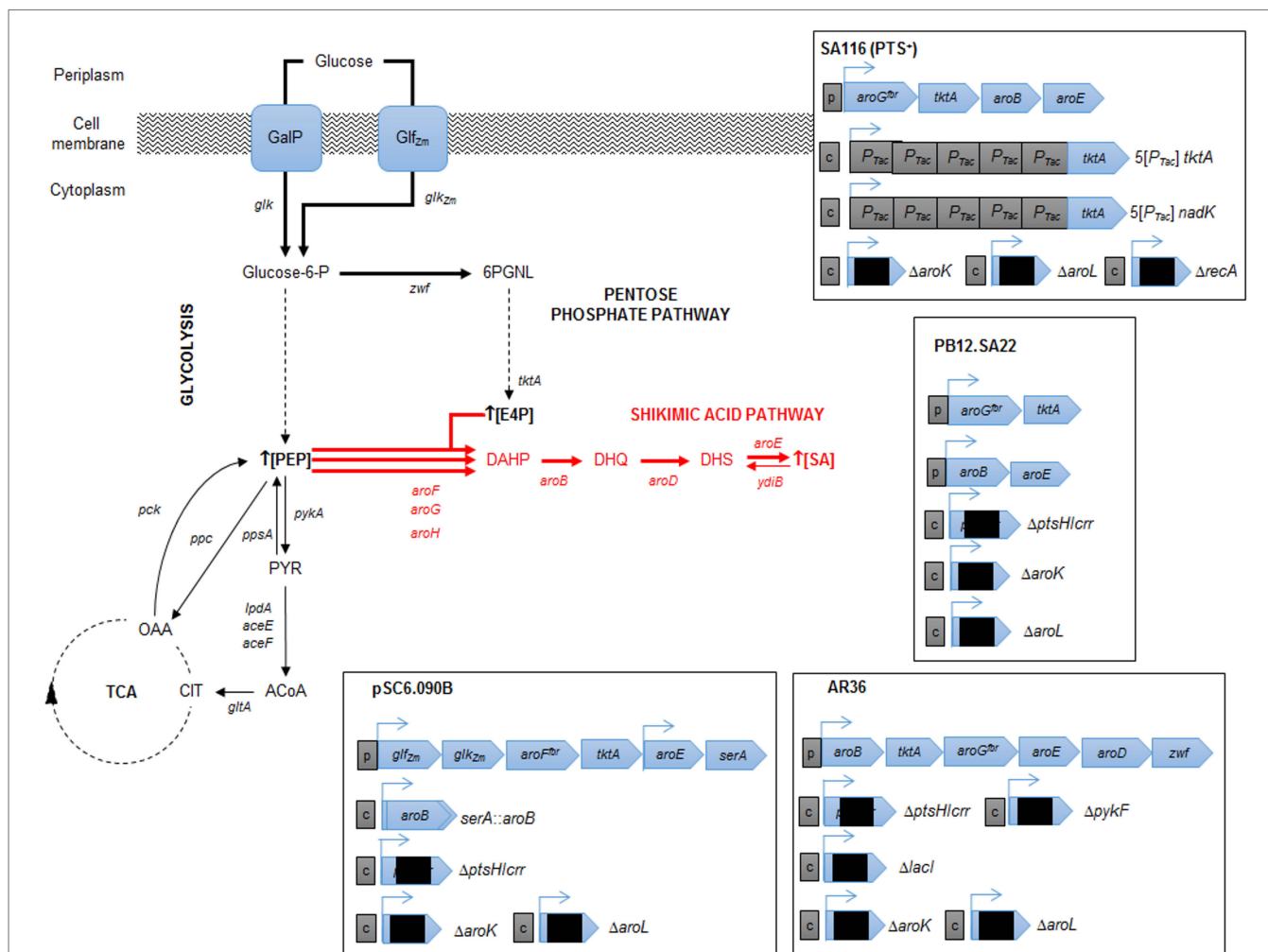
glc, glucose; YE, yeast extract.

PEP pool was to eliminate the PTS system, which would not only modify the amount of PEP but also redistribute the stoichiometric matrix to raise the maximum theoretical yield to 86% (mol/mol) (Chandran et al., 2003; Krämer et al., 2003). The main problem with this solution is the resultant low level of glucose transport, which results in a strain with hampered growth (PTS<sup>-</sup> phenotype) (Flores et al., 1996, 2007; Aguilar et al., 2012). Nevertheless, various strategies have been developed to revert this low glucose consumption and low growth phenotype. Using rational ME strategies, substitution of the PTS for another glucose transport system has been performed. Chandran et al. used heterologous expression of the *Zymomonas mobilis* (Glf) glucose transporter, a glf-encoded glucose facilitator and a glk-encoded glucose kinase (Glk), thereby allowing cells to consume glucose more efficiently without consuming PEP (Frost et al., 2002; Chandran et al., 2003; Krämer et al., 2003). Another strategy is to apply laboratory adaptive evolution onto a PTS<sup>-</sup> strain. Flores et al. used a continuous culture with glucose as a single carbon source to select high glucose consumption-evolved derivative strains (PTS<sup>-</sup> glc<sup>+</sup> phenotype). Characterization of these mutants revealed overexpression of the *galP* and *glk* genes encoding galactose permease and glucokinase, respectively, allowing and improving glucose transport and phosphorylation capabilities and resulting in an increased specific growth rate and PEP availability (Flores et al., 1996, 2007; Aguilar et al., 2012).

Finally, with precursors known to induce redirection, deregulation, and overexpression of the SA pathway genes *aroB*, *aroD*, and *aroE* have been achieved, resulting in an efficient carbon flux

from CCM to the SA pathway. The highest production to date corresponds to the SP1.1pts-/pSC6.090B strain, a PTS<sup>-</sup> derivative strain with a plasmid containing two *tac* promoters, the first of which controls expression of the *glf*, *glk*, *aroF<sup>thr</sup>*, and *tktA* genes and the second of which controls expression of the *aroE* and *serA* genes (Chandran et al., 2003). The reasoning behind this construction was to increase the PEP pool by deleting PTS, to recuperate glucose consumption by overexpressing *glf* and *glk*, to assure E4P pool enhancement by overexpressing *tktA* and to induce a deregulated pull toward the AA pathway via an *aroF<sup>thr</sup>*, as discussed before. The second promoter in the plasmid was designed to overexpress *aroE*, allowing continuous flux of the SA pathway; additionally, a second copy of *aroB* was introduced into the chromosome instead of reintroducing the serine production-related gene *serA* to the cell via the plasmid and was used as a selection marker for plasmid retention. This approximation, along with the deletion of genes related to SA consumption (*aroK* and *aroL*), allowed SA accumulation, achieving a production capacity of 87 g/L SA, with a yield of 36% (mol/mol) and a productivity of ~5.3 g/L h when a 10 L glucose-fed batch was cultured. This strain has the highest titer accumulation recorded in the literature to date (Table 1; Figure 3).

Even with these rational strategies, the yield of the SP1.1pts-/pSC6.090B strain is far from the theoretical maximum yields of PTS<sup>-</sup> derivative strains. In 2010, Escalante et al. presented a JM101 PTS<sup>-</sup> derivative strain with high glucose consumption capacity that was capable of overexpressing the *galP* and *glk* genes and that was produced from an adaptive evolution process. This



**FIGURE 3 | Relevant engineered *E. coli* strains for SA production.** Metabolic traits of *E. coli* derivative strains engineered for SA production resulting in highest SA titer and yield from glucose. The figure illustrates alternative glucose transporter GalP (galactose permease) selected by the cell after laboratory adaptive evolution process of a PTS<sup>-</sup> mutant (Flores et al., 1996, 2007; Aguilar et al., 2012). Glf (glucose facilitator) and Glk (glucokinase) from *Z. mobilis* (plasmid cloned). Resultant characteristics of engineered strains are shown for pSC6.090B (Chandran et al., 2003), PB12.SA22 (Escalante et al., 2010), AR36 (Rodríguez et al., 2013), and SA116 strain (Cui et al., 2014). CCM key intermediates and protein encoding genes: TCA, tricarboxylic acid pathway; E4P, erythrose-4-P; 6PGNL, 6-phospho-D-glucono-1,5-lactone; PEP, phosphoenolpyruvate; PYR, pyruvate; ACoA, acetyl-CoA; CIT, citrate; OAA, oxaloacetate; *zwf*, glucose 6-phosphate-1-dehydrogenase; *tktA*, transketolase I; *pykA*, pyruvate kinase II; *lpdA*, *aceE* and *aceF*, coding for PYR dehydrogenase subunits; *glfA*, citrate synthase; *pck*, PEP carboxylase; *ppsA*, PEP synthetase. SA pathway intermediates and genes: DAHP, 3-deoxy-D-arabino-heptulosonate-7-phosphate; DHQ, 3-dehydroquinate; DHS, 3-dehydroshikimate; SA, shikimic acid. Continuous arrows indicate single enzymatic reactions; dashed arrows show several enzymatic reactions. Bold arrows show improved carbon flux. Black squares in plasmids/operons indicate gene interruption; c, chromosomal gene interruption or integration; p, plasmid-cloned genes.

strain (PB12), along with a two-plasmid expression system for *aroG<sup>lbr</sup>-tktA* and *aroB-aroE*, respectively, under *lacUV5* promoters inducible by IPTG (PB12.SA22), allowed a yield of 29% (mol/mol) (Escalante et al., 2010). Further modifications allowed them to find that a *pykF* deletion could result in higher yields of total aromatic compounds, up to 50% (mol/mol), even when presenting an SA yield diminution (0.21%). In this case, the amount of flux reduced from PEP to PYR was redirected throughout the SA pathway, and without the correct amounts of enzymes, new bottlenecks appeared, causing other metabolites and intermediates to accumulate (Escalante et al., 2010). Therefore, it was clear that regulating gene expression and dosage remained a problem for more efficiently redirecting flux not only toward but also within

the SA pathway. Regarding that topic, Rodríguez et al. utilized the PB12 *pykF-aroKL*<sup>-</sup> strain and developed a plasmid with a constitutively strong promoter onto a synthetic operon containing the *aroB*, *tktA*, *aroG<sup>lbr</sup>*, *aroE*, *aroD*, and *zwf* genes (AR36) for synchronous expression of the relevant genes found in previous research. With this expression design, the AR36 derivative strain is able to redirect the carbon flow to SA even in high glucose conditions (above 100 g/L of the initial substrate concentration) without producing high acetate titers. This strain produced up to 43 g/L of SA via simple batch processes, with SA yields of 42% (mol/mol) and total SA pathway intermediate yield up to 67% of the theoretical maximum, representing the highest yield managed to be produced to date (Rodríguez et al., 2013) (Table 1; Figure 3).

Regarding the expression and regulation of key SA production genes, most of the research has been performed using plasmid expression; however, there are multiple drawbacks, ranging from structural and segregational instability to metabolic burden, of plasmid replication. Cui et al. (2014) resolved this problem by constructing a strain with an *aroG<sup>fb</sup>*, *aroB*, *aroE*, and *tktA* gene cluster integrated into the chromosome and by tuning the copy number and expression by using chemically induced chromosomal evolution (CICHe) with triclosan. They also overexpressed the *ppsA* and *csrB* genes to enhance the PEP pyruvate pool. This strain rendered a 1.70 g/L SA titer, with a yield up to 0.25 (mol/mol). Finally, they studied and improved cofactor availability for SA production optimization; in this case, NADPH availability was increased because *aroE*-encoded enzymes require this specific cofactor for the DHS to SA conversion reaction. By plasmid-based or chromosomal overexpression of the NADPH availability-related genes *pntAB* or *nadK*, this cofactor pool was enhanced, which was directly correlated to the SA production capabilities of the strain. As they changed the promoters and the expression of all the chromosomally inserted genes related to SA production mentioned above, they managed to construct a strain capable of producing a yield of 0.33 (mol/mol) SA from glucose (Figure 3).

Many other examples of SA production platforms in *E. coli* have been studied in the literature, the most relevant of which are referred to in Table 1, rendering many industrially competitive strains and processes. Nevertheless, the main efforts throughout the past two decades were directed toward a particular pathway approach. As shown in Table 1, few SA production processes have been designed utilizing an overview of global regulation and manipulation, which can be obtained from omics data. Transforming this global information into global knowledge on the complexity of cell regulation would reveal the existing regulatory bottlenecks, allowing us to metabolically engineer potential strains using a systems biology approach, finally ensuring a truly rational strain design with optimized production capabilities.

## Omics Approaches for the Study of the SA Pathway in *Escherichia coli*

Classical ME approaches applied to diverse *E. coli* strains to obtain SA-overproducing derivatives have targeted key genes in the CCM and SA pathways, allowing successful reconfiguration of the biochemical network of engineered strains and resulting in the efficient redirection of carbon flow from CCM to SA production. However, the inactivation of key genes coding for enzymes involved in global regulatory processes, such as the PTS system or coding for key node enzymes, such as the PykF enzyme results in global metabolic reconfiguration, which frequently introduces significant flux imbalances. This often produces undesirable outcomes, including the accumulation of intermediates, feedback inhibition of upstream enzymes, the formation of unwanted byproducts, and the diminution of cellular fitness via the rerouting of resources toward the unnecessary or non-essential production of pathway enzymes. By understanding these newly created flux imbalances in SA-overproducing derivative strains,

it is possible to boost the overall cellular physiology, product titer, productivity, and yield, taking into account a global view of cellular metabolism (Biggs et al., 2014). Combinatorial approaches allow researchers to work with this scenario by conducting global cellular searches, but the necessity for high-throughput screening is often a drawback for pathway engineering. The other approach is to augment knowledge and computational tools to properly predict designs to achieve a desired metabolic outcome (Fong, 2014). Several high-throughput approximations, such as genomic, transcriptomic, and proteomic predictions, have been applied to aromatic AAs and engineered SA-overproducing strains for the identification of non-intuitive targets other than those genes/enzymes involved in the CCM and SA pathways that might be suitable for further modification by ME.

## The Identification of YdiB (*ydiB*) as a Key Enzyme in Byproduct Formation During SA Synthesis

The analysis of available genome sequences using Hidden Markov Model profiles to identify all known enzymes of the SA pathway has shown that some genes have been lost in diverse microbial groups, particularly in host-associated bacteria (Zucko et al., 2010). This condition has been proposed to result in the development of undesirable metabolic traits, such as the hydroaromatic equilibration observed in *E. coli*, resulting in the synthesis of so-called missing metabolites, such as quinic acid (QA) and DHQ, by a reversion of the SA biosynthetic pathway (Knop et al., 2001; Zucko et al., 2010). The coproduction of high quantities of the byproducts DHS and QA is not a desirable trait; they significantly reduce the SA yield because QA is co-purified during the downstream process of SA purification from the culture supernatant (Knop et al., 2001; Krämer et al., 2003; Diaz Quiroz et al., 2014).

The strain W3110.shik1 ( $\Delta$ *aroL*, *aroG<sup>fb</sup>*, *trpE<sup>fb</sup>*, and *tnaA*) engineered for SA production growing in low glucose (high phosphate) or glucose-rich (low phosphate) conditions resulted in the production of SA in cultures with mineral broth, as the single inactivation of shikimate kinase II (*aroL*) allows carbon flux to CHA through shikimate kinase I (*aroK*), resulting in the synthesis of aromatic AAs. However, under carbon-limited conditions, SA production decreased by 59%, and the byproducts DHS, DHQ, gallic acid (GA), and QA were detected in the culture supernatant with respect to phosphate limiting culture conditions (Johansson et al., 2005). Global transcriptomic analysis (GTA) of the strain W3110.shik1 in chemostatic culture conditions, comparing between glucose and phosphate limiting conditions, allowed identification of the significantly upregulated genes *ydiB* (coding for shikimate dehydrogenase/quinic acid dehydrogenase), *aroD*, and *ydiN*, which encodes a putative transporter, in carbon limiting conditions. The upregulation of these genes, particularly *ydiB* (10× with respect to its paralogs, *aroE*), was proposed to increase the YdiB level, which uses DHQ and SA as substrates, as this enzyme has a lower  $K_m$  for SA in the presence of NAD<sup>+</sup> (Keseler et al., 2013). Additionally, the intracellular concentration of NAD<sup>+</sup> is reported to be 40-fold higher than that of NADH<sup>+</sup>, suggesting that the dehydrogenase activity on SA to produce DHS is favored by YdiB *in vivo* (Johansson and



Lidén, 2006). These results suggest that byproduct formation during SA production was associated with the reversal of the biosynthetic pathway from (1)  $\text{SA} + \text{NAD(P)}^+ \leftrightarrow \text{DHS} + \text{NAD(P)}\text{H} + \text{H}^+$  and (2)  $\text{DHS} + \text{NAD(P)}\text{H} + \text{H}^+ \leftrightarrow \text{QA} + \text{NAD(P)}^+$  by YdiB or (3)  $\text{DHS} + \text{H}_2\text{O} \leftrightarrow \text{DHQ}$  by AroD (Figure 4). The presence of a large amount of intracellular SA was proposed to drive the reversal of the pathway, whereas YdiN was proposed to be the exporter of the aromatic byproducts (Johansson and Lidén, 2006).

As these results suggest an important role of YdiB in byproduct synthesis during SA production and its intracellular accumulation under glucose limiting conditions, a rational strategy to avoid byproduct synthesis was the inactivation of *ydiB* and/or the upregulation of its paralogs, *aroE*, coupled to efficient SA secretion from the cell. The upregulation of *aroE* expression (simultaneously with other key genes of the CCM and SA pathways) in  $\text{PTS}^- \text{ gluc}^+ \text{ aroK}^- \text{ aroL}^-$  engineered strains resulted in the highest SA titer and yield reported with low byproduct formation (Chandran et al., 2003; Rodríguez et al., 2013) (Table 1).

The replacement of *ydiB* by its paralogs, *aroE*, in a modular biosynthetic pathway design for L-tyrosine production in *E. coli* MG1655 resulted in the elimination of a bottleneck caused by the high affinity of YdiB protein for the accumulation of QA and DHS. This replacement in the modular plasmid construction  $\text{P}_{lac-UV5} \text{ aroE, aroD, aroB}^{op}, \text{ aroG}^{tr}, \text{ ppsA, tktA}$  (*op* = optimize codon usage) resulted in the accumulation of 700 mg/L of SA, which was in turn successfully channeled to L-tyrosine (Juminaga et al., 2012). However, combinational plasmid overexpression of the *aroB, aroD, aroE, ydiB, aroK, aroL, aroA, aroC*, and *tyrB* genes with *ydiB* resulted in high L-tyrosine production. This result suggested that *ydiB* but not its paralog, *aroE*, is an attractive target for the overproduction of this aromatic AA because *aroE* in *E. coli* codes for a feedback-inhibited shikimate dehydrogenase, resulting

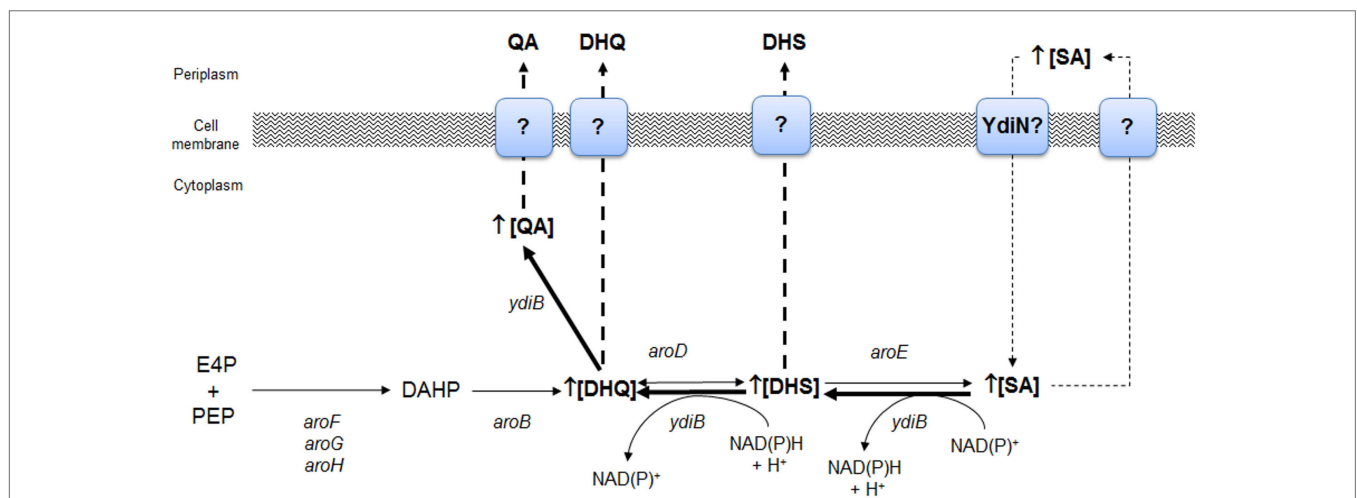
in a bottleneck for L-tyrosine production (Lütke-Eversloh and Stephanopoulos, 2008).

### The Impact of *pykF* Inactivation on the Protein Levels of SA Pathway Enzymes

The pyruvate kinase isoenzymes Pyk I and Pyk II (coded by *pykF* and *pykA*, respectively) play key roles in CCM via Pyk activity, together with 6-phospho-fructokinase I (coded by *pkfA*) and glucokinase (*glk*), controlling carbon flux through the glycolytic pathway (Keseler et al., 2013). Pyk I and Pyk II are key allosteric enzymes that catalyze one of the two substrate-level phosphorylation steps yielding ATP and the irreversible trans-phosphorylation of PEP and ADP into PYR and ATP, maintaining a permanent flux of PYR to acetyl-CoA (Keseler et al., 2013).

Inactivation of the *pykF* gene in *E. coli*  $\text{PTS}^-$  derivatives (PB12 strain) engineered for SA production has resulted in the increased flux of carbon into the SA pathway (Escalante et al., 2010), increasing the DAHP concentration above 370% (and the total SA pathway aromatic yield) with respect to the *pykF*<sup>+</sup> parental strain. Further applications of ME strategies in the PB12 strain *pykF*<sup>-</sup> resulted in the derivative strain AR36, which produces up to 40 g/L SA with a yield of 0.42 mol SA/mol glc (Table 1) (Rodríguez et al., 2013), demonstrating that the inactivation of *pykF* in a  $\text{PTS}^-$  derivative strain significantly improves PEP flux toward SA synthesis.

Global proteomic analysis in a *pykF*<sup>-</sup> derivative of *E. coli* (BW25113) compared with its *pykF*<sup>+</sup> parental strain revealed the differential overexpression of 24 proteins, including enzymes from the SA pathway and aromatic AAs. The upregulation of key SA pathway enzymes, including the DAHPS AroG isoenzyme (2.66 times more abundant with respect to the *pykF*<sup>+</sup> strain), which is involved in the synthesis of DAHP, the first intermediate of the SA pathway, and the AroB enzyme (DHQ synthase, 4.72



**FIGURE 4 | Identification of key genes of the SA pathway involved in the biosynthesis of aromatic byproducts QA and DHS from SA as determined by global transcriptomic analysis in *E. coli* W3110.shik1.** Overexpression of *ydiB*, *aroD*, and *ydiN* genes allowed proposing that under carbon limiting growth conditions, SA is intracellularly accumulated as consequence of an inefficient export to periplasmic space or as consequence of its back transport to the cytoplasm as consequence of extracellular accumulation. YdiN, a putative transporter coded by *ydiN* was proposed to be involved in SA back import. Backflow of SA to DHS was possibly catalyzed by YdiB, whereas synthesis of DHQ from DHS was performed by AroD enzyme and finally, YdiB performed synthesis of QA from DHQ. Adapted from Johansson and Lidén (2006).

times more abundant with respect to the *pykF*<sup>+</sup> strain) (Kedar et al., 2007). These results support the positive impact of *pykF*<sup>-</sup> inactivation not only on increased PEP availability but also on increased carbon flux toward the SA pathway.

### The Identification of Other Possible Key Catabolic and Biosynthetic Genes Involved in SA Production

Batch fermentation cultures of the *E. coli* PB12.SA22-derivative strain for SA production (PTS<sup>-</sup> Glc<sup>+</sup>aroK<sup>-</sup>, *aroL*<sup>-</sup>aroG<sup>thr</sup>, *tktA*, *aroB*, and *aroD*; **Table 1**) using complex production media containing 25 g/L glucose and 15 g/L yeast extract (YE) showed two characteristic growth stages: a fast growth phase associated with low glucose consumption during the first 8–10 h of cultivation and low SA production, and a second slow growth stage with high glucose consumption until this carbon source was completely consumed (25 h of cultivation). Interestingly, SA production continues during the STA phase after glucose, used as a carbon source, was completely consumed, until the end of fermentation (50 h) (Escalante et al., 2010). This behavior suggested that during the EXP growth phase, this strain preferentially consumed some YE components to support growth, whereas glucose was used to produce SA and other pathway intermediates, suggesting the existence of regulatory and physiological differences between EXP and STA phases (Cortés-Tolalpa et al., 2014).

GTA was performed to corroborate this hypothesis during SA production in batch fermentation cultures using complex fermentation broth (Chandran et al., 2003; Escalante et al., 2010; Rodriguez et al., 2013) by comparing global expression profiling between the mid-exponential growth phase (EXP, 5 h of cultivation), the early stationary phase (STA1, 9 h) and the late STA phase (44 h); EXP/STA1, EXP/STA2, and STA1/STA2 comparisons were conducted (Cortés-Tolalpa et al., 2014) (**Figure 5**).

The relevant results showed EXP growth in the derivative strain PB12.SA22 during the first 9 h of cultivation. When the L-tryptophan provided by YE available in the supernatant was completely consumed (6 h), the strain entered the low-growth phase (even in the presence of glucose) until 26 h of cultivation, when glucose was completely consumed; this was associated with low SA production. Interestingly, during the stationary stage, SA production continued until the end of fermentation (50 h), achieving the highest accumulation (7.63 g/L of SA) in the absence of glucose (**Figures 5A,B**, upper panel).

GTA comparisons among EXP/STA1, EXP/STA2, and STA1/STA2 showed no significant differences in the regulation of genes from the CCM and SA pathways, but for the EXP/STA1 comparison, the upregulation of genes coding for sugar transport, AA catabolism and biosynthesis, and nucleotide/nucleoside salvage was observed (**Figure 5A**). Interestingly, in the STA2 phase, the highest SA production was observed in the absence of glucose in supernatant, associated with the upregulation of genes encoding transporters for the AAs L-lysine, L-arginine, L-histidine, L-ornithine, and L-glutamic acid and enzymes involved in the synthesis, interconversion, and catabolism of L-arginine. As all of these AAs are provided by YE, this result suggests that this AA could play a key role in fueling carbon to SA synthesis, and likely also in L-arginine conversion to the

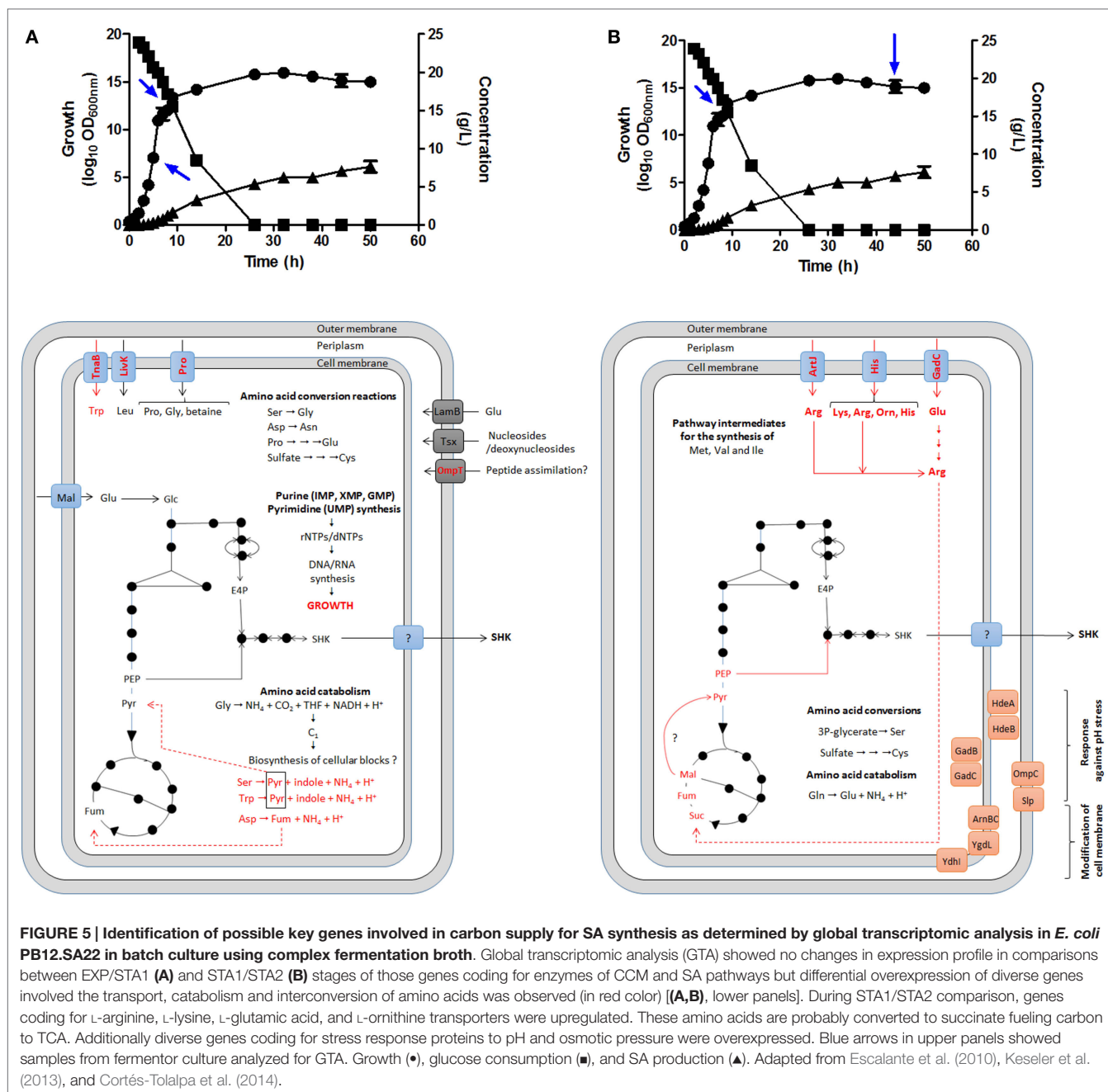
TCA intermediate succinate through the super-pathway of L-arginine and L-ornithine degradation (Keseler et al., 2013) (**Figure 5B**). These results indicate the origin of carbon required for the highest SA production during the STA phase after glucose was completely consumed. Additionally, the upregulation of genes involved in the pH stress response and inner and outer membrane modifications suggests a cellular response to environmental conditions imposed on the cell at the end of fermentation (44 h) (Cortés-Tolalpa et al., 2014).

The upregulation of genes coding for the biosynthesis and interconversion pathways of almost all AAs was also observed by GTA in cultures under C-limiting condition of the derivative strain W3110.shik1 grown in minimal broth. These changes were postulated to correlate to aromatic AA starvation with these culture conditions, although this strain maintained functional shikimate kinase I (*aroK*), allowing the accumulation of SA but maintaining carbon flux toward CHA and aromatic AAs (Johansson and Lidén, 2006).

As demonstrated by GTA in the SA-producing strain PB12.SA22 during batch culture fermentations in complex media containing YE, several metabolic constraints limit the growth capabilities of this strain, stopping growing even in the presence of glucose. The highest SA production observed in the late stationary stage in the absence of glucose was probably supported by the non-aromatic AA content of YE. This evidence supports valuable information to further optimize culture strategies, as YE feeding increased the SA titer and yield in engineered strains.

### Omics Data Integration into Metabolic Modeling: Moving Toward Data Integration for Rational Strain Improvement

Although ME is capable of reconfiguring a biochemical network to redirect the substrate conversion into valuable compounds by manipulating the microorganism genetic code, its classical rational approach often introduces significant new flux imbalances. This has often caused undesirable outcomes due to the accumulation of intermediates, feedback inhibition of upstream enzymes, and the formation of unwanted byproducts of cellular fitness diminution via the rerouting of resources toward the unnecessary or non-essential production of pathway enzymes (Biggs et al., 2014). By understanding these newly created flux imbalances on mutant strains, it is possible to boost overall cellular health and the product titer, productivity, and yield, taking into account a holistic view of cellular metabolism (Biggs et al., 2014). Since the development of the omics, there has been an increased interest to understand the behavior of complete biological systems. Omics renders biological data from all levels of metabolism going all the way from genome to metabolome, these data combined give us the possibility to study the whole organism instead of single components. To achieve this, mathematical models play the important role of converting omics data into organismal information and knowledge (Åkesson et al., 2004; Fong, 2014). There are several frameworks and approaches for the mathematical modeling of metabolism developed to collect high-throughput data to understand as well as to predict phenotypic



function. Computational applications have been developed using models as quantitative mathematical representations of biological systems and or their components to a suitable level of simplification (Jouhten, 2012). These computational tools can be used to identify new biological pathways in the host microorganism for the selection and improvement of important genotypic characteristics to improve the production of the desired compound (Long et al., 2015). In this section, we discuss some mathematical models and computational tools that can be used in ME to utilize all high-throughput omics data and render new insights into flux distributions, regulation constraints, and modification targets to optimize the production of desired metabolites.

To understand the challenges and virtues of mathematical modeling, we must observe that biological systems are complex in nature, involving the transport of information through many layers, including the genome, transcriptome, proteome, and metabolome; therefore, regulatory steps between the interactions of these layers finally render the complex outcome of the phenotypic behavior (Cloots and Marchal, 2011; Fong, 2014). Therefore, mathematical models have been used to evolve and clarify the complex network interactions and system characteristics to reveal the underlying mechanisms. Despite this high degree of complexity, with all the recent advances and data sets available, mathematical modeling promises to generate experimentally

testable hypotheses, predictions, and new insights into systems biology to better understand cell behavior (Stelling, 2004).

The first step in mathematical modeling is reconstructing the metabolic network. With the advent of the genomic era since approximately 1999, reconstruction can be achieved on a genome-wide scale for many organisms and has been used to expand the knowledge on metabolic networks and to identify new or non-intuitive metabolic reactions to be engineered for further strain improvements (Åkesson et al., 2004; Kim et al., 2012). Genome-scale models are assembled and manually curated from the annotated genome, and biochemical information is used to render a representation of the metabolic network on which mathematical representations will set a matrix of equations to model its behavior. The reconstruction of a genomic metabolic network starts through the examination and identification of the coding regions or open reading frames on the sequence. After analysis with established algorithms and biochemical and physiological databases (EcoCyc, MPW, and KEGG WIT), sequences can be converted into feasible reactions, and a metabolic network can be reconstructed from genomic information (Covert et al., 2001). This reconstructed network, based on genomic data, is now the backbone of an *in silico* organism. Many organisms have been completely sequenced and have simultaneously been extensively biochemically studied, which in turn can make the reconstructed metabolic network more complete (Covert et al., 2001). In recent years, ~40% of all eukaryotic models and 30% of the total prokaryotic models have been published, advancing from highly characterized organisms (*E. coli* and *Saccharomyces cerevisiae*) to less characterized species with more complex biological systems that have special characteristics for specific applications (Kim et al., 2012). When a network is described with sufficient detail, some qualitative predictions can be made, and with the inclusion of stoichiometric, thermodynamic, and kinetic data, the reconstructed metabolic map of an organism can be used to generate quantitative predictions regarding phenotype via the construction of mathematical models (Covert et al., 2001). For example, individual genes have been deleted from *in silico* models, and correlations between the model and experimental data for the consequences of each deletion have been found to be 60% accurate for *Helicobacter pylori* and 86% accurate for *E. coli* (Price et al., 2003). Nevertheless, the challenges for the construction of these *in silico* models include obtaining high-throughput data to reconstruct more complete models, which can be sorted out by using omics data and combinatorial experimentation, and constructing mathematical approaches to model and render specific solutions for the highly complex systems of biological networks. Because genome-scale metabolic networks comprise hundreds to thousands of reactions, a large number of parameters are required to mathematically describe networks, which, therefore, requires the development of informatic intensive modeling approaches to describe its complexity and to make useful predictions regarding phenotypic behavior for strain design (Price et al., 2003).

The most used approaches are those arising from stoichiometric modeling, which uses mass balances over the metabolic network and assumes a pseudo-steady-state condition to determine intracellular metabolic fluxes, along with additional experimental data to solve the underdetermined linear equation

system (Åkesson et al., 2004). Stoichiometric modeling creates a matrix ( $S$ ) for the metabolites and metabolic reactions, in which each element indicates a stoichiometric coefficient, along with a vector that contains all of the unknown reaction rates ( $v$ ); under the steady state assumption, flux distribution will be represented by  $S \cdot v = 0$  (Jouhten, 2012; Kim et al., 2012). As expected, this equation system will have many solutions, or more precisely, it will render a convex solution space, and because genome-scale metabolic models include all possible metabolic reactions whether or not they are expressed, meaningful solutions must be narrowed down to render a viable solution (Kim et al., 2012). The main problem is that due to the high number of equations and parameters, these systems are always underdetermined; thus, the use of thermodynamic, metabolic, kinetic, and all other experimental data available is required to impose constraints, to reveal a plausible solution, and therefore to conduct quantitative analysis and make predictions regarding cell behavior (Fong, 2014).

To accomplish this desirable outcome, mathematical modeling researchers have developed many approaches to render the complexity, including the use of interaction-based, constraint-based, and mechanism-based methodologies for calculations. Interaction-based approaches isolate autonomous units performing distinct functions in cellular systems, accounting for modularity, which simplifies networks and systems to perform a topological analysis to reveal the principles of cellular organization. Constraint-based approaches account for the physicochemical invariance of networks in addition to network topology. This approach along with stoichiometric modeling, is capable of confining the numerous steady-state flux distributions the metabolic reconstruction network can have (convex space of solutions) into a smaller group, which complies with the constraints indicated by the knowledge regarding the system (a set of feasible states). Even so, this approach accounts only for the steady state, and therefore produces static models; thus, the final phenotypic behavior in changing intracellular or extracellular environments is difficult to address (Stelling, 2004). Mechanism-based approaches use kinetic parameters along with stoichiometric parameters to render the dynamic behavior of cells; thus, such approaches can formulate precise flux distributions and explore the regulation over time. The main problem with this approach is that the knowledge on mechanisms and associated parameters (kinetic reaction parameters) has, thus, far been limited, as so much effort and so many resources must be used to accomplish this type of models (Stelling, 2004; Jouhten, 2012; Long et al., 2015).

Constraint-based approaches are the most used ones to date because of their capability to render flux distribution modeling even with a relatively small amount of information. These approaches state the constraints under which the reconstructed network operates based on stoichiometry and thermodynamics, including directionality and biochemical loops (Price et al., 2003). Such constraints can be imposed by linear optimization; for example, standard flux-base analysis (FBA) uses growth optimization, selecting only the flux solutions, that in turn, produce the maximum growth rate for network topology (Åkesson et al., 2004). Newer flux solution reduction methods have been developed to study the solution space, accounting for the



optimization of not only growth but also many other linear and non-linear objective functions, such as the maximum biomass, maximum ATP, minimum overall intracellular flux, maximum ATP yield per flux unit, maximum biomass yield per flux unit, maximum substrate consumption, minimum number of reaction steps, minimum redox potential, and minimum flux production between others (Price et al., 2003; Schuetz et al., 2007). These optimization principles, along with other constraints arising from specific conditions being either biotic (e.g., the inactivation, subexpression, or overexpression of specific target genes) or abiotic (e.g., aerobic culture, anaerobic culture, nitrogen limitation, carbon limitation, available substrates), will help not only to render the most feasible flux distribution solution but also to study the consequences of changing the genetic cellular output or fermentation parameters for a specific objective. This information is of great use for ME because it renders the ability through different modeling frameworks to study and predict the effects of knocking out genes, tuning the expression of target genes involved in specific reactions, network robustness, the endpoint of adaptive evolution, the identification and characterization of regulation, and heterologous reactions and *de novo* reactions on strain design. There are many reviews that discuss and compare multiple modeling frameworks, such as OptGene, OptStrain, CosMos, OptFocrcce, FaceCon, and FOCAL, for the constraint-based analysis of genome-wide metabolic networks (Price et al., 2003; Schuetz et al., 2007; Krull and Wittmann, 2010; Cloots and Marchal, 2011; Jouhten, 2012; Fong, 2014; King et al., 2015; Long et al., 2015). In this review, we will focus only on one or two framework examples given the scope of this work.

The first strain design method involving knockouts is OptKnock, a bi-level optimization framework used to identify optimal reaction deletion strategies, coupling cellular growth, and target metabolite production. OptKnock identifies deletions with the highest chemical production within the solution space obtained by the maximum growth rate constraint (Long et al., 2015). Pharkya et al. (2003) used this framework to explore the overproduction of amino acids; specifically for AA, they addressed the channeling of flux from PEP to AA by removing the *ppc* gene, which could lead to the redirection of carbon flux to the formation of CHA via the accompanying deletions of pyruvate oxidase, pyruvate dehydrogenase, and pyruvate lyase reactions. The deletion of *ppc* by itself fails to redirect PEP to AA; the ability to detect its contribution though the co-inactivation of other reactions is a very useful tool of ME because the classical experimental deletion of this gene would have produced negative results for pathway optimization. In other words, *in silico* modeling enables researchers to avoid designs toward a local maxima or minima when trying to identify the modifications required to achieve a global maxima for their specific purposes.

FBA with grouping reaction constraints (FBAwGR) was developed to improve the accuracy of metabolic simulation by incorporating the grouping of reaction constraints of functionally and physically related reactions in the model. This framework allows the consideration of genomic context and flux-converging analyses. Genomic context accounts for conserved neighborhoods, gene fusion, and co-occurrences of genes to organize fluxes that are likely to be on or off together. Flux-converging

analyses then restrict the carbon flux solution space to the number of metabolites participating in reactions and converging patterns from a specific carbon source. This framework has been used to predict changes in flux patterns caused by several genetic modifications, such as *pykF*, *zwf*, *ppc*, and *sucA* deletions in *E. coli*, showing good agreements with experimentally obtained fluxes (Kim et al., 2012).

Regarding the scope of this review for SA production in *E. coli*, we have found few studies in the literature that account for metabolic modeling. Nevertheless, the notable work by Chen et al. (2011), described FBA constraint analysis by stoichiometry and mass balance, assuming no growth and optimizing SA as the objective function to design modifications for the production of intermediate metabolites of the aromatic pathway. The model identified several key reaction steps for overexpression, similarly to those previously reported for AA optimization (overexpression of the *aroF*, *tktA*, *ppsA*, and *glf* genes, as well as deletions of the *ldhA* and *ackA* genes) by avoiding carbon waste through lactate and acetate fluxes. Finally, with all of the modifications made, their model identified the *zwf* gene as the critical node for redirection of the carbon flux into the AA pathway; its deletion led to an optimized accumulation of QA, GA, and SA, accounting for a 47% molar conversion of glucose (Chen et al., 2014).

Regarding other SA related work, Rizk and Liao (2009), managed to use EM to model, study, and predict DAHP production in *E. coli* toward aromatic production. Ensemble modeling (EM) is a mechanism-based modeling approach that decomposes metabolic reactions into elementary reaction steps, incorporating all available phenotypic observations for the wild type and mutant strains, integrating this information into the mathematical approach to identify the kinetic variables of each elementary reaction step (Rizk and Liao, 2009; Khodayari et al., 2014). Rizk and Liao (2009), using different flux bounds on the pathway split ratio between glycolysis and the PPP. Then, by using data from literature for overexpression of the *tktA*, *talA*, and *pps* genes, they were able to screen the solution space models compared with the phenotypic behavior, selecting the ones that properly described the experimental data (from a 1500 solution space to 7, 171, and 195 solution spaces, according to glycolysis:PPP ratios of 25:75, 75:25, and 95:5, respectively). This subset of flux solutions revealed that TktA is the first controlling rate step and that PPS, only with simultaneous overexpression of TktA can augment DAHP production; these findings are in accordance with the phenotypic observations in the literature. Based on these results, they conclude that the flux distributions found could be reverse engineered to enhance aromatic production in *E. coli* (Rizk and Liao, 2009).

Notably, despite the existence of many genome-scale metabolic models and various mathematical approaches, many of the fluxes remain undetermined, as many solutions remain plausible. Thus, more information is needed to ensure the modeling quality by the validation and incorporation of *in vivo* experimental data. These experimental data can be acquired from transcriptomic, proteomic, or fluxomic data. Strategies incorporating these extensive experimental data have been developed to enhance the quality and the accuracy of metabolic models (Kim et al., 2012). Fluxomic data in its core provide us with the most important

information as fluxes are the modeling outcome, but experimental procedures can only be used for relatively smaller networks and in specific conditions. Nevertheless, these data are of utmost importance and are commonly used to validate model solutions to flux distributions. ME models have been used to integrate protein expression data to reconstruct and add constraints to genome-level metabolic models, relating kinetic equations into catalytic constraints to approximate stoichiometric relationships between enzyme abundance and catalyzed fluxes (O'Brien and Palsson, 2015). This integration of proteomic data adds thermodynamic and allocation constraints that help in the identification of a consistent flux state, allowing an explanation of aspects of cell behavior and relationships that have remained elusive, such as the interaction of ribosomes with metabolism, carbon limited to carbon excess metabolic shifts, substrate uptake regulation, membrane protein relationships, and other protein spatial constraints that can utterly dominate and/or change metabolic responses (O'Brien and Palsson, 2015). Transcriptomic data have been used to exploit the regulatory information in the expression data to provide additional constraints for the metabolic fluxes in the model by analyzing if or when gene expression correlates with a given metabolic flux (Åkesson et al., 2004; Kern et al., 2006). Computational protocols have been developed for this type of data integration, such as mixed integer linear programming (MILP), which seeks to maximize the agreement between experimental data and computational fluxes by limiting the presentation of entities with the capability to carry flux; meanwhile, the flux of absent entities would be 0 (Fong, 2014). Åkesson et al. (2004) used gene expression microarray data from chemostat and batch cultures of *S. cerevisiae* to create Boolean variables for all of the reactions encompassed on a genome-scale metabolic model to ascertain the absent/present fluxes using analysis software. These new constraints allowed the computation of metabolic flux distributions to enhance the metabolic behavior in batch cultures, along with the quantitative prediction of exchange fluxes as well as the qualitative estimation of changes in intracellular fluxes compared with the model without transcription constraints, as verified by experimental measurements of flux (Åkesson et al., 2004).

Many methods have been developed to introduce transcriptomic regulation into modeling predictions, such as probabilistic regulation of metabolism (PROM), which calculates the probability that a metabolic target gene will be expressed relative to the activity of its regulating transcription factor, metabolic adjustment by differential expression (MADE), which creates a sequence of binary expression states so that when the gene expression changes from one condition to another, the flux reaction will change in accordance with its value, and gene inactivity moderated by metabolism and expression (GIMME), which is a context metabolic model that predicts the subsets of reactions used under a particular condition using gene expression data and which identifies a flux distribution to optimize a given biological objective, such as growth and/or ATP production, along with FBA (Kim et al., 2012). Finally, a method called E-Flux can map continuous gene expression into flux bound constraints according to gene-protein-reaction (GPR) associations, limiting the upper and lower bounds on fluxes so that genes expressed at higher

levels will result in higher flux values (Kim et al., 2012). This and other methods have been reviewed and compared by Machado and Herrgård (2014), who concluded that the prediction of flux levels from gene expression remains far from solved because the predictions obtained by simple FBA with growth maximization and parsimony criteria were as good or even better than those obtained using the incorporation of transcriptomic data. Nevertheless, they acknowledge that some methods evaluated give reasonable predictions under certain conditions that there is no universal method that performs well under all scenarios and that the transcriptome should provide some guidelines for the correct phenotype determination within the space of solutions resulting from the large number of degrees of freedom in metabolic networks, recommending that users should perform a careful evaluation of the meaningfulness of the results for their particular applications (Machado and Herrgård, 2014).

There are many successful mathematical modeling approaches to produce good and accurate predictions of phenotypic behavior in the literature; all of these methods help us to understand and simplify metabolic regulation and systems to comprehend and find new or non-intuitive targets for ME. Even so, there is still much work to be conducted to understand and construct better models of metabolic networks. There are many challenges because cell behavior is a complex system that, therefore, has complex outcomes and regulation. These challenges range from network reconstruction, mathematical treatments, and true flux distribution determination to the integration of all systems data (omics) to achieve regulation and phenotypic predictions. Nevertheless, the effort put into understanding this matter has produced and will continue to produce new insights for strain design and ME. Explaining all the considerations, challenges and achievements in this field is not within the scope of this review as many reviews have been published on these matters (Liu et al., 2004; Patil et al., 2004; Stelling, 2004; Schuetz et al., 2007; Kim et al., 2012; Machado and Herrgård, 2014; Saha et al., 2014; Long et al., 2015; O'Brien and Palsson, 2015). Rather, this review is aimed to provide the reader with interesting findings and perspective on how the mathematical modeling of biological systems can be and is useful for ME, especially regarding SA and AA production, for which these methods can be of relevance to exploit the maximum production capability of *E. coli* that remains unachieved.

## Summary and Perspectives

SA is a key intermediate of the common aromatic pathway with diverse applications in the synthesis of valuable pharmaceutical compounds, but major interest relies on SA as the precursor for the chemical synthesis of OSF, the neuraminidase inhibitor of diverse influenza viruses, including pandemic strains. Diverse efforts have been made to produce high titers and yields of SA in metabolically engineered strains of *E. coli* with successful genetic modifications, including the following: (1) interruption of the SA pathway by the inactivation of shikimate kinase coding genes (*aroK* and *aroL*), which results in the high accumulation of SA; (2) increasing the intracellular availability of the CCM intermediate PEP by inactivation of the PTS system and replacing this glucose translocation system by other housekeeping or

heterologous glucose transporters and by inactivation of the *pykF* gene; and (3) the overexpression of diverse key genes of the CCM and SA pathways, such as *zwf*, *tktA*, *aroB*, *aroD*, and *aroE*, under the control of constitutively expressed or inducible promoters in plasmid-cloned operons or chromosome-integrated copies. These engineered strains have been cultured in batch or fed-batch culture conditions using a complex fermentation media including glucose and YE, resulting in the highest titer and yield of SA reported (Chandran et al., 2003; Rodriguez et al., 2013).

The above-described genetic changes impose global nutritional, regulatory, and metabolic constraints on the resultant engineered strains, which must be explored to determine their relevance on SA production. GTA of the SA-producing strain W3110.shik1 provided evidence supporting the roles of *ydiB*-, *aroD*-, and *ydiN*-encoded proteins in byproduct formation during SA production under glucose limiting conditions (Johansson and Lidén, 2006). Recent ME strategies applied for L-tyrosine (Juminaga et al., 2012) and SA production (Rodriguez et al., 2013) demonstrated the relevance of *ydiB* inactivation and *aroD* overexpression to avoid byproduct formation and to improve carbon flux toward the desired aromatic products.

Interruption of the SA pathway by inactivation of the *aroK* and *aroL* genes imposes an auxotrophic requirement for aromatic AAs and probably other metabolites derived from CHA on the cell; these effects were successfully reversed by the addition of YE to the fermentation media.

As the chemical complexity of YE or peptone significantly interferes in the study of carbon flux through the CCM and SA pathway metabolic networks, no studies to date have been reported on the application of metabolic models to identify possible targets for the application of further ME strategies focused on the improvement of SA production in fermentation culture using complex production media (Chandran et al., 2003; Escalante et al., 2010; Chen et al., 2012; Rodriguez et al., 2013; Cui et al., 2014). The application of omics, such as GTA, in SA-producing conditions, including YE, as reported for the strain P12.SA22, provides valuable information on the role

of diverse transporter systems and other pathways involved in carbon supply from YE to SA synthesis (Cortés-Tolalpa et al., 2014). These results highlight the relevance of information retrieved from the application of omics, such as GTA, or proteomic approaches in successful aromatic compound-producing strains to obtain data for mathematical modeling of metabolism.

Further application of synthetic biology strategies based on modular combinational design including key genes from the CCM and SA pathways in operons and optimized codon usage, and the construction of continuous genetic modules regulated by the same promoter but coupled to an efficient translational level by the selection of efficient ribosome binding sites (RBS) from tailored-made RBS libraries are promising strategies for the subsequent optimization of SA-producing strains. These synthetic strategies have been applied for the efficient production of L-tyrosine in *E. coli* (Juminaga et al., 2012) and for the successful production of SA in *Corynebacterium glutamicum* (Zhang et al., 2015), respectively. Great advances in SA production in *E. coli* have been made over the past decades. However, more and new developments must be made, taking into account the vast, recently acquired data from omics technology. These data, along with their integration with ME technology and experience, can lead to more global insight into cell physiology, allowing new engineering techniques from a systems ME perspective to be identified and developed.

## Author Contributions

All authors participated equally in the preparation of this contribution. All authors have read and approved the final manuscript.

## Acknowledgments

This work was supported by CONACYT Ciencia Básica project 240519.

## References

- Aguilar, C., Escalante, A., Flores, N., de Anda, R., Riveros-McKay, F., Gosset, G., et al. (2012). Genetic changes during a laboratory adaptive evolution process that allowed fast growth in glucose to an *Escherichia coli* strain lacking the major glucose transport system. *BMC Genomics* 13:385. doi:10.1186/1471-2164-13-385
- Åkesson, M., Förster, J., and Nielsen, J. (2004). Integration of gene expression data into genome-scale metabolic models. *Metab. Eng.* 6, 285–293. doi:10.1016/j.ymben.2003.12.002
- Báez, J. L., Bolívar, F., and Gosset, G. (2001). Determination of 3-deoxy-D-arabino-heptulosonate 7-phosphate productivity and yield from glucose in *Escherichia coli* devoid of the glucose phosphotransferase transport system. *Biotechnol. Bioeng.* 73, 530–535. doi:10.1002/bit.1088
- Báez-Viveros, J. L., Osuna, J., Hernández-Chávez, G., Soberón, X., Bolívar, F., and Gosset, G. (2004). Metabolic engineering and protein directed evolution increase the yield of L-phenylalanine synthesized from glucose in *Escherichia coli*. *Biotechnol. Bioeng.* 87, 516–524. doi:10.1002/bit.20159
- Biggs, B. W., De Paeppe, B., Santos, C. N. S., De Mey, M., and Kumaran Ajikumar, P. (2014). Multivariate modular metabolic engineering for pathway and strain optimization. *Curr. Opin. Biotechnol.* 29, 156–162. doi:10.1016/j.copbio.2014.05.005
- Bongaerts, J. (2001). Metabolic engineering for microbial production of aromatic amino acids and derived compounds. *Metab. Eng.* 3, 289–300. doi:10.1006/mben.2001.0196
- Chandran, S. S., Yi, J., Draths, K. M., von Daeniken, R., Weber, W., and Frost, J. W. (2003). Phosphoenolpyruvate availability and the biosynthesis of shikimic acid. *Biotechnol. Prog.* 19, 808–814. doi:10.1021/bp025769p
- Chen, K., Dou, J., Tang, S., Yang, Y., Wang, H., Fang, H., et al. (2012). Deletion of the *aroK* gene is essential for high shikimic acid accumulation through the shikimate pathway in *E. coli*. *Bioresour. Technol.* 119, 141–147. doi:10.1016/j.biortech.2012.05.100
- Chen, P. T., Chiang, C.-J., Wang, J.-Y., Lee, M.-Z., and Chao, Y.-P. (2011). Genomic engineering of *Escherichia coli* for production of intermediate metabolites in the aromatic pathway. *J. Taiwan Inst. Chem. Eng.* 42, 34–40. doi:10.1016/j.jtice.2010.03.010
- Chen, X., Li, M., Zhou, L., Shen, W., Algasan, G., Fan, Y., et al. (2014). Metabolic engineering of *Escherichia coli* for improving shikimate synthesis from glucose. *Bioresour. Technol.* 166, 64–71. doi:10.1016/j.biortech.2014.05.035
- Chen, X., Zhou, L., Tian, K., Kumar, A., Singh, S., Prior, B. A., et al. (2013). Metabolic engineering of *Escherichia coli*: a sustainable industrial platform for bio-based chemical production. *Biotechnol. Adv.* 31, 1200–1223. doi:10.1016/j.biotechadv.2013.02.009
- Cloots, L., and Marchal, K. (2011). Network-based functional modeling of genomics, transcriptomics and metabolism in bacteria. *Curr. Opin. Microbiol.* 14, 599–607. doi:10.1016/j.mib.2011.09.003



- Cortés-Tolalpa, L., Gutiérrez-Ríos, R. M., Martínez, L. M., De Anda, R., Gosset, G., Bolívar, F., et al. (2014). Global transcriptomic analysis of an engineered *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system during shikimic acid production in rich culture medium. *Microb. Cell Fact.* 13, 28. doi:10.1186/1475-2859-13-28
- Covert, M. W., Schilling, C. H., Famili, I., Edwards, J. S., Goryanin, I. I., Selkov, E., et al. (2001). Metabolic modeling of microbial strains *in silico*. *Trends Biochem. Sci.* 26, 179–186. doi:10.1016/S0968-0004(00)01754-0
- Cui, Y.-Y., Ling, C., Zhang, Y.-Y., Huang, J., and Liu, J.-Z. (2014). Production of shikimic acid from *Escherichia coli* through chemically inducible chromosomal evolution and cofactor metabolic engineering. *Microb. Cell Fact.* 13, 21. doi:10.1186/1475-2859-13-21
- Díaz Quiroz, D. C., Carmona, S. B., Bolívar, F., and Escalante, A. (2014). Current perspectives on applications of shikimic acid aminoshikimic acids in pharmaceutical chemistry. *Res. Rep. Med. Chem.* 4, 35–46. doi:10.2147/RRMC.S46560
- Draths, K. M., Knop, D. R., and Frost, J. W. (1999). Shikimic acid and quinic acid: replacing isolation from plant sources with recombinant microbial biocatalysis. *J. Am. Chem. Soc.* 121, 1603–1604. doi:10.1021/ja9830243
- Escalante, A., Calderón, R., Valdivia, A., de Anda, R., Hernández, G., Ramírez, O. T., et al. (2010). Metabolic engineering for the production of shikimic acid in an evolved *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system. *Microb. Cell Fact.* 9, 21. doi:10.1186/1475-2859-9-21
- Escalante, A., Salinas Cervantes, A., Gosset, G., and Bolívar, F. (2012). Current knowledge of the *Escherichia coli* phosphoenolpyruvate-carbohydrate phosphotransferase system: peculiarities of regulation and impact on growth and product formation. *Appl. Microbiol. Biotechnol.* 94, 1483–1494. doi:10.1007/s00253-012-4101-5
- Estevez, A., and Estevez, R. (2012). A short overview on the medicinal chemistry of (-)-shikimic acid. *Mini Rev. Med. Chem.* 12, 1443–1454. doi:10.2174/138955712803832735
- F. Hoffmann-La Roche Ltd. (2009). *Roche Annual Report 09*. Basel, Switzerland: F. Hoffmann-La Roche Ltd. 127 p. Available at: <http://www.roche.com/gb09e.pdf>
- Flores, N., Leal, L., Sigala, J. C., de Anda, R., Escalante, A., Martínez, A., et al. (2007). Growth recovery on glucose under aerobic conditions of an *Escherichia coli* strain carrying a phosphoenolpyruvate: carbohydrate phosphotransferase system deletion by inactivating *arcA* and overexpressing the genes coding for glucokinase and galactose permease. *J. Mol. Microbiol. Biotechnol.* 13, 105–116. doi:10.1159/000103602
- Flores, N., Xiao, J., Berry, A., Bolívar, F., and Valle, F. (1996). Pathway engineering for the production of aromatic compounds in *Escherichia coli*. *Nat. Biotechnol.* 14, 620–623. doi:10.1038/nbt0596-620
- Fong, S. S. (2014). Computational approaches to metabolic engineering utilizing systems biology and synthetic biology. *Comput. Struct. Biotechnol. J.* 11, 28–34. doi:10.1016/j.csbj.2014.08.005
- Frost, J. W., Frost, K. M., and Knop, D. R. (2002). *Biocatalytic Synthesis of Shikimic Acid*. Available at: <http://www.google.com/patents/US6472169> (accessed July 5, 2015).
- Ghosh, S., Chisti, Y., and Banerjee, U. C. (2012). Production of shikimic acid. *Biotechnol. Adv.* 30, 1425–1431. doi:10.1016/j.biotechadv.2012.03.001
- Gosset, G. (2009). Production of aromatic compounds in bacteria. *Curr. Opin. Biotechnol.* 20, 651–658. doi:10.1016/j.copbio.2009.09.012
- Johansson, L., and Lidén, G. (2006). Transcriptome analysis of a shikimic acid producing strain of *Escherichia coli* W3110 grown under carbon- and phosphate-limited conditions. *J. Biotechnol.* 126, 528–545. doi:10.1016/j.jbiotec.2006.05.007
- Johansson, L., Lindskog, A., Silfversparre, G., Cimdander, C., Nielsen, K. F., and Lidén, G. (2005). Shikimic acid production by a modified strain of *E. coli* (W3110.shik1) under phosphate-limited and carbon-limited conditions. *Biotechnol. Bioeng.* 92, 541–552. doi:10.1002/bit.20546
- Jouhten, P. (2012). Metabolic modelling in the development of cell factories by synthetic biology. *Comput. Struct. Biotechnol. J.* 3, 1–9. doi:10.5936/csbj.201210009
- Juminaga, D., Baidoo, E. E. K., Redding-Johanson, A. M., Bath, T. S., Burd, H., Mukhopadhyay, A., et al. (2012). Modular engineering of L-tyrosine production in *Escherichia coli*. *Appl. Environ. Microbiol.* 78, 89–98. doi:10.1128/AEM.06017-11
- Kedar, P., Colah, R., and Shimizu, K. (2007). Proteomic investigation on the *pyk-F* gene knockout *Escherichia coli* for aromatic amino acid production. *Enzyme Micro. Technol.* 41, 455–465. doi:10.1016/j.enzmictec.2007.03.018
- Kern, R., Malki, A., Abdallah, J., Tagourt, J., and Richarme, G. (2006). *Escherichia coli* HdeB is an acid stress chaperone. *J. Bacteriol.* 189, 603–610. doi:10.1128/JB.01522-06
- Keseler, I. M., Mackie, A., Peralta-Gil, M., Santos-Zavaleta, A., Gama-Castro, S., Bonavides-Martínez, C., et al. (2013). EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res.* 41, D605–D612. doi:10.1093/nar/gks1027
- Khodayari, A., Zomorodi, A. R., Liao, J. C., and Maranas, C. D. (2014). A kinetic model of *Escherichia coli* core metabolism satisfying multiple sets of mutant flux data. *Metab. Eng.* 25, 50–62. doi:10.1016/j.ymben.2014.05.014
- Kim, T. Y., Sohn, S. B., Kim, Y. B., Kim, W. J., and Lee, S. Y. (2012). Recent advances in reconstruction and applications of genome-scale metabolic models. *Curr. Opin. Biotechnol.* 23, 617–623. doi:10.1016/j.copbio.2011.10.007
- King, Z. A., Lloyd, C. J., Feist, A. M., and Palsson, B. O. (2015). Next-generation genome-scale models for metabolic engineering. *Curr. Opin. Biotechnol.* 35, 23–29. doi:10.1016/j.copbio.2014.12.016
- Knop, D. R., Draths, K. M., Chandran, S. S., Barker, J. L., von Daeniken, R., Weber, W., et al. (2001). Hydroaromatic equilibration during biosynthesis of shikimic acid. *J. Am. Chem. Soc.* 123, 10173–10182. doi:10.1021/ja0109444
- Krämer, M., Bongaerts, J., Bovenberg, R., Kremer, S., Müller, U., Orf, S., et al. (2003). Metabolic engineering for microbial production of shikimic acid. *Metab. Eng.* 5, 277–283. doi:10.1016/j.ymben.2003.09.001
- Krull, R., and Wittmann, C. (2010). *Linking Cellular Networks and Bioprocesses*. Heidelberg: Springer.
- Li, S., Yuan, W., Wang, P., Zhang, Z., Zhang, W., and Ownby, S. (2007). *Method for the Extraction and Purification of Shikimic Acid*. Available at: <http://www.google.com/patents/US20070149805>
- Lin, S., Liang, R., Meng, X., OuYang, H., Yan, H., Wang, Y., et al. (2014). Construction and expression of mutagenesis strain of *aroG* gene from *Escherichia coli* K-12. *Int. J. Biol. Macromol.* 68, 173–177. doi:10.1016/j.ijbiomac.2014.04.034
- Liu, A., Liu, Z. Z., Zou, Z. M., Chen, S. Z., Xu, L. Z., and Yang, S. L. (2004). Synthesis of (+)-zeilenone from shikimic acid. *Tetrahedron* 60, 3689–3694. doi:10.1016/j.tet.2004.02.066
- Liu, Q., Cheng, Y., Xie, X., Xu, Q., and Chen, N. (2012). Modification of tryptophan transport system and its impact on production of L-tryptophan in *Escherichia coli*. *Bioresour. Technol.* 114, 549–554. doi:10.1016/j.biortech.2012.02.088
- Long, M. R., Ong, W. K., and Reed, J. L. (2015). Computational methods in metabolic engineering for strain design. *Curr. Opin. Biotechnol.* 34, 135–141. doi:10.1016/j.copbio.2014.12.019
- Lütke-Eversloh, T., and Stephanopoulos, G. (2008). Combinatorial pathway analysis for improved L-tyrosine production in *Escherichia coli*: identification of enzymatic bottlenecks by systematic gene overexpression. *Metab. Eng.* 10, 69–77. doi:10.1016/j.ymben.2007.12.001
- Machado, D., and Herrgård, M. (2014). Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism. *PLoS Comput. Biol.* 10:e1003580. doi:10.1371/journal.pcbi.1003580
- Matsuoka, Y., and Shimizu, K. (2012). Importance of understanding the main metabolic regulation in response to the specific pathway mutation for metabolic engineering of *Escherichia coli*. *Comput. Struct. Biotechnol. J.* 3, 1–10. doi:10.5936/csbj.201210018
- O'Brien, E. J., and Palsson, B. O. (2015). Computing the functional proteome: recent progress and future prospects for genome-scale models. *Curr. Opin. Biotechnol.* 34, 125–134. doi:10.1016/j.copbio.2014.12.017
- Patil, K. R., Åkesson, M., and Nielsen, J. (2004). Use of genome-scale microbial models for metabolic engineering. *Curr. Opin. Biotechnol.* 15, 64–69. doi:10.1016/j.copbio.2003.11.003
- Patnaik, R., and Liao, J. C. (1994). Engineering of *Escherichia coli* central metabolism for aromatic metabolite production with near theoretical yield. *Appl. Environ. Microbiol.* 60, 3903–3908.
- Patnaik, R., Spitzer, R. G., and Liao, J. C. (1995). Pathway engineering for production of aromatics in *Escherichia coli*: confirmation of stoichiometric analysis by independent modulation of AroG, TktA, and Pps activities. *Biotechnol. Bioeng.* 46, 361–370. doi:10.1002/bit.260460409
- Pharkya, P., Burgard, A. P., and Maranas, C. D. (2003). Exploring the overproduction of amino acids using the bilevel optimization framework OptKnock. *Biotechnol. Bioeng.* 84, 887–899. doi:10.1002/bit.10857
- Price, N. D., Papin, J. A., Schilling, C. H., and Palsson, B. O. (2003). Genome-scale microbial *in silico* models: the constraints-based approach. *Trends Biotechnol.* 21, 162–169. doi:10.1016/S0167-7799(03)00030-1
- Raghavendra, T. R., Vaidyanathan, P., Swathi, H. K., Ramesha, B. T., Ravikanth, G., Ganeshiah, K. N., et al. (2009). Prospecting for alternate sources of shikimic acid, a precursor of Tamiflu, a bird-flu drug. *Curr. Sci.* 96, 771–772.



- Rangachari, S., Friedamna, T. C., Hartmann, R., and Weisenfeld, R. B. (2013). *Processes for Producing and Recovering Shikimic Acid*. Available at: <http://www.google.com/patents/US8344178> (accessed July 5, 2015).
- Rizk, M. L., and Liao, J. C. (2009). Ensemble modeling for aromatic production in *Escherichia coli*. *PLoS ONE* 4:e6903. doi:10.1371/journal.pone.0006903
- Rodríguez, A., Martínez, J. A., Báez-Viveros, J. L., Flores, N., Hernández-Chávez, G., Ramírez, O. T., et al. (2013). Constitutive expression of selected genes from the pentose phosphate and aromatic pathways increases the shikimic acid yield in high-glucose batch cultures of an *Escherichia coli* strain lacking PTS and *pykF*. *Microb. Cell Fact.* 12, 86. doi:10.1186/1475-2859-12-86
- Rodríguez, A., Martínez, J. A., Flores, N., Escalante, A., Gosset, G., and Bolívar, F. (2014). Engineering *Escherichia coli* to overproduce aromatic amino acids and derived compounds. *Microb. Cell Fact.* 13, 126. doi:10.1186/s12934-014-0126-z
- Saha, R., Chowdhury, A., and Maranas, C. D. (2014). Recent advances in the reconstruction of metabolic models and integration of omics data. *Curr. Opin. Biotechnol.* 29, 39–45. doi:10.1016/j.copbio.2014.02.011
- Scheiwiller, T., and Hirschi, S. (2010). *09 Roche Annual Report*. Finance Report. Basel: F. Hoffmann-La Roche Ltd.
- Schuetz, R., Kuepfer, L., and Sauer, U. (2007). Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol. Syst. Biol.* 3, 119. doi:10.1038/msb4100162
- Stelling, J. (2004). Mathematical models in microbial systems biology. *Curr. Opin. Microbiol.* 7, 513–518. doi:10.1016/j.mib.2004.08.004
- Wang, G.-W., Hu, W.-T., Huang, B.-K., and Qin, L.-P. (2011). *Illicium verum*: a review on its botany, traditional use, chemistry and pharmacology. *J. Ethnopharmacol.* 136, 10–20. doi:10.1016/j.jep.2011.04.051
- World Health Organization. (2011). *Pandemic Influenza Preparedness. Framework for the Sharing of Influenza Viruses and Access to Vaccines and Other Benefits*. France: World Health Organization. Available at: [http://www.who.int/influenza/resources/pip\\_framework/en/](http://www.who.int/influenza/resources/pip_framework/en/)
- Yi, J., Li, K., Draths, K. M., and Frost, J. W. (2002). Modulation of phosphoenolpyruvate synthase expression increases shikimate pathway product yields in *E. coli*. *Biotechnol. Prog.* 18, 1141–1148. doi:10.1021/bp020101w
- Zhang, B., Zhou, N., Liu, Y.-M., Liu, C., Lou, C.-B., Jiang, C.-Y., et al. (2015). Ribosome binding site libraries and pathway modules for shikimic acid synthesis with *Corynebacterium glutamicum*. *Microb. Cell Fact.* 14, 71. doi:10.1186/s12934-015-0254-0
- Zucko, J., Dunlap, W. C., Shick, J. M., Cullum, J., Cercelet, F., Amin, B., et al. (2010). Global genome analysis of the shikimic acid pathway reveals greater gene loss in host-associated than in free-living bacteria. *BMC Genomics* 11:628. doi:10.1186/1471-2164-11-628

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Martínez, Bolívar and Escalante. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Fluorescent reporter libraries as useful tools for optimizing microbial cell factories: a review of the current methods and applications

*Frank Delvigne\*, H  l  ne P  cheux and C  dric Tarayre*

*Microbial Processes and Interactions (MiPI), Gembloux Agro-Bio Tech, University of Liège, Gembloux, Belgium*

## OPEN ACCESS

**Edited by:**

Hilal Taymaz Nikerel,  
Boğazici University, Turkey

**Reviewed by:**

Marjan De Mey,  
Ghent University, Belgium  
Lothar Eggeling,  
Forschungszentrum Jülich, Germany

**\*Correspondence:**

Frank Delvigne,  
Microbial Processes and Interactions  
(MIPI), Gembloux Agro-Bio Tech,  
University of Liège,  
Passage des Déportés 2,  
Gembloux 5030, Belgium  
f.delvigne@ulg.ac.be

**Specialty section:**

*This article was submitted to Systems Biology, a section of the journal Frontiers in Bioengineering and Biotechnology*

Received: 30 June 2015

**Accepted:** 11 September 2015

**Published:** 28 September 2015

**Citation:**

Delvigne F, Pêcheux H and Tarayre C  
(2015) Fluorescent reporter libraries  
as useful tools for optimizing  
microbial cell factories: a review of  
the current methods and  
applications.  
*Front. Bioeng. Biotechnol.* 3:147.  
doi: 10.3389/fbioe.2015.00147

The use of genetically encoded fluorescent reporters allows speeding up the initial optimization steps of microbial bioprocesses. These reporters can be used for determining the expression level of a particular promoter, not only the synthesis of a specific protein but also the content of intracellular metabolites. The level of protein/metabolite is thus proportional to a fluorescence signal. By this way, mean expression profiles of protein/metabolites can be determined non-invasively at a high-throughput rate, allowing the rapid identification of the best producers. Actually, different kinds of reporter systems are available, as well as specific cultivation devices allowing the on-line recording of the fluorescent signal. Cell-to-cell variability is another important phenomenon that can be integrated into the screening procedures for the selection of more efficient microbial cell factories.

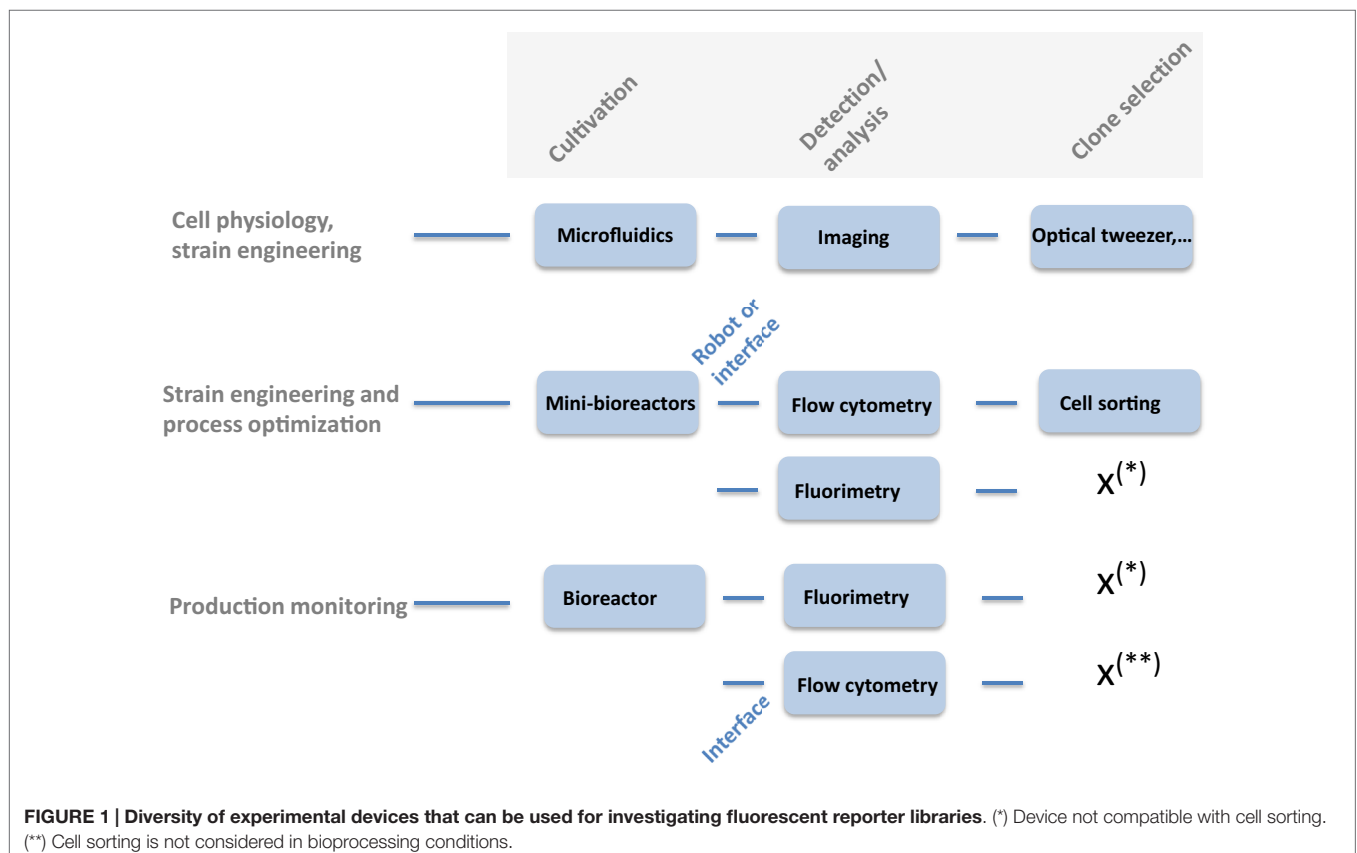
**Keywords:** mini-bioreactors, high-throughput, flow cytometry, single cell

## Fluorescent Proteins as Reporter: From the Single-Cell Challenge to the Advent of Synthetic Biology

Besides its importance in giving new insights in cell biology, e.g., for the analysis of the intrinsic and extrinsic component of phenotypic noise among microbial population (Swain et al., 2002), fluorescent reporters are also an important component for the development of new bioprocesses, i.e., for strain engineering and process optimization up to the large-scale production of bioproducts (Polizzi and Kontoravdi, 2014). From a fundamental perspective, fluorescent reporter libraries are available for several model organisms, including *Escherichia coli* K12 MG1655 (Zaslaver et al., 2006) and *Saccharomyces cerevisiae* (Newman et al., 2006). The two above-mentioned fluorescent reporter libraries have notably been used for the characterization of noise in protein expression (Newman et al., 2006; Silander et al., 2012). Indeed, molecular processes associated with DNA transcription and translation are subjected to different noise mechanisms leading a cell-to-cell variability in protein content among an isogenic microbial population (Sanchez et al., 2013). Clone libraries and experimental devices for the cultivation and the detection of fluorescent signal at the single-cell level have been specifically developed (Taniguchi et al., 2010). Besides these genome-scale libraries, fluorescent reporter system can also be used for the design of smaller libraries, e.g., for the estimation of the strength of several promoters that could be used for the expression of a protein of interest or for the design of synthetic metabolic pathways (Xu et al., 2012). This last application of fluorescent reporter is very important, since synthetic biology becomes widespread for the design

of efficient cell factories, able to synthesize fuels and chemicals with high titer. Fluorescent proteins can also be found in more specific applications, such as the detection of the intracellular metabolite level (Schallmey et al., 2014) or the control of lab evolution (Reyes, 2012a,b), which will be detailed throughout this review. The exploitation of a fluorescent reporter library is greatly facilitated by the use of specific experimental devices. Indeed, the actual experimental toolbox dedicated to the use of fluorescent reporters allows for all the manipulations required in bioprocess optimization and scale-up and comprise specific cultivation devices, analytical tools, and clone selection tools (**Figure 1**). Among the cultivation device, a full range of culture volume is available, from micro- (picoliter) and mini-bioreactor (milliliter) to full-scale bioreactors (liter). Micro-bioreactors are based on microfluidic chips adapted to the culture of microorganisms. A nice example of micro-bioreactor has been developed by Grunberger et al. (2012, 2013), where a single microbial cell is isolated in a picoliter chamber perfused by fresh medium. The height of the picoliter chamber is designed in order to be slightly higher than the mean diameter of the microbial cells, so that microbial cells are maintained in the chamber and are continually fed with fresh medium whereas metabolites and by-products are continuously extracted. The perfusion mode of culture allows thus to cultivate microorganisms under constant environmental conditions. Imaging allows for the acquisition of the individual division rate and also the gene activity if linked with a fluorescent reporter system. A major limitation of the

actual micro-bioreactor is that they are not designed to work in the operating modes generally met in industrial conditions, i.e., batch and fed-batch (Love, 2013; Grunberger et al., 2014). This limitation can be overcome by considering mini-bioreactors. This range of bioreactor involves the use of cultivation volume of around 1 ml (Klockner and Buchs, 2012). One of the most advanced mini-bioreactor platform to date is the Biolector system, and its extension Robolector (Funke et al., 2010). This device is based on a microplate and allows the parallel cultivation of 48 samples with on-line determination of biomass, pH, dissolved oxygen, and fluorescence. High oxygen transfer efficiency allows to carry out microbial culture in fully aerobic conditions and fed-batch and pH control are available, ensuring the compatibility of the results with those gained in conventional stirred bioreactor. Fluorescence sensor available in each well can be used to gain informations at the level of a fluorescent reporter system, but only at the bulk level. Other mini-bioreactor systems are now available; either based on the concept of “shaken” bioreactor (e.g., Micro 24-microreactor system developed by Pall) or “stirred” bioreactor (e.g., the 48-bioreactor system developed by 2mag) (Lattermann and Buchs, 2014). Single-cell results can be obtained by coupling the cultivation device to a robotic platform delivering the samples to a flow cytometer. Microbial phenotypic heterogeneity is a phenomenon that has gained a lot of attention, considering its potential impact on bioprocesses (Delvigne et al., 2014). Fluorescent reporter library is a technology of choice for investigating the effect of microbial phenotypic heterogeneity



on bioprocesses and single-cell analytical devices compatible with bioreactor are needed at this level. Indeed, genome-scale investigation of GFP reporter libraries have led to a better understanding of the evolution of noise in gene and protein expression (Taniguchi et al., 2010; Silander et al., 2012) and the associated molecular mechanisms (Swain et al., 2002). Since on-line flow cytometry can be adapted to monitor phenotypic heterogeneity during a microbial culture by coupling a specific interface to the cultivation device (Brognaux et al., 2013), these mechanisms are now considered in bioprocessing conditions (Polizzi and Kontoravdi, 2014; Baert et al., 2015). Several interfaces have been built for this purpose and have been successfully used to monitor the activity of fluorescent reporter in bioprocess conditions (Abu-Absi et al., 2003; Arnoldini et al., 2013; Besmer et al., 2014; Delvigne et al., 2015). In the following section, the use of fluorescent reporter libraries will be illustrated for different biotechnological applications.

## Biotechnological Applications Related to the Use of Fluorescent Reporter Systems

### Production of Recombinant Proteins: From Promoter Strength to Noise in Expression

The use of transcriptional reporter, i.e., a promoter linked to the sequence of a fluorescent protein, allows for the fast and non-invasive characterization of the strength of an expression vector and the protein synthesis rate over time (DeLisa, 1999; Zaslaver et al., 2004). This type of information is very useful when optimizing a process and is greatly facilitated by the fluorescent signal that can be detected on-line (see Figure 1 for an overview of the combination between cultivation devices and recording tools available). Translational reporters, i.e., a promoter linked to the sequence of a protein of interest either homologous or heterologous, this latter being tagged with a fluorescent protein facilitating the detection of the whole chimeric protein, can be used for the optimization of the production of a recombinant protein, i.e., by adjusting the induction profile if the promoter is inducible or more generally by adjusting the cultivation conditions (DeLisa, 1999; Abu-Absi et al., 2003). However, several factors must be taken into account for the design of an efficient translational reporter. Indeed, using a fluorescent protein as a tag can have a profound impact on protein folding and stability, and specific technologies, such as split-GFP, have been designed for this purpose (Cabantous et al., 2005). Another drawback is that the signal gained from a fluorescent reporter accounts only for a product accumulated intracellular. It is thus theoretically not possible to get informations about the amount of products secreted, which is an important limitation since most of the biotechnological applications are oriented toward product secretion to the extracellular medium in order to decrease the costs associated with downstream processing operations. In this context, a specific micro-cultivation device based on micro-engraving has been developed and allows not only for the analysis of the intracellular level of fluorescent protein but also for the amount of fluorescent protein excreted at the extracellular level (Love, 2010, 2012). Another strategy for the detection

of excreted compounds include the use of surface display or droplet microfluidics (Hai and Magdassi, 2004; Kintsjes et al., 2012; Mazutis et al., 2013).

### Detection of Metabolite Productivity at the Intracellular Level

Besides the use of promoter-based fluorescent reporter, other biosensing strategies can be designed for the detection of intracellular metabolites. The most widely used strategy for the detection of intracellular metabolites relies on the use of transcriptional factors (Eggeling et al., 2015). In this context, the interaction of the metabolite with a transcriptional factor induces the synthesis of the fluorescent reporter molecule. Another strategy relies on the use of riboswitches where the binding of the metabolite to a RNA aptamer triggers the synthesis of the fluorescent reporter molecule. However, this strategy is less used than the design based on transcriptional factors. This kind of sensing strategy is invaluable for the design of hyper productive strains (Eggeling et al., 2015). For example, Binder and co-workers used transcription factor-based biosensor for the selection of L-lysine hyperproducers among a population of *Corynebacterium glutamicum* exposed to chemical mutagenesis (Binder, 2012, 2013). The fluorescent biosensor facilitated the high-throughput detection by flow cytometry, as well as the isolation of the mutants by fluorescence-activated cell sorting (FACS). A similar workflow has also been applied to other relevant cell factories, such as *S. cerevisiae* and *E. coli* (Delvigne and Goffin, 2014).

### Synthetic Biology and the Optimization of Artificial Metabolic Pathways

Synthetic biology allows for more important modification of cell factories, such as the integration of artificial metabolic pathways on specific microbial chassis (Silva-Rocha and de Lorenzo, 2010; Martinez-Garcia et al., 2014). However, synthetic biology requires the use of reliable parts for the design of effective cell factories. Such parts can now be found in dedicated libraries, such as those of the BioBrick initiative (Rokke et al., 2014). However, the biobrick registry is based mainly on standardized parts designed for easily assembling synthetic vectors. Since this initiative, other registries have been designed on the basis of more functional parts. As an example, the ePathBrick library can be used to insert artificial metabolic pathways inside *E. coli* for the synthesis of diverse chemicals (Xu et al., 2012). Such libraries and assembly protocols allow for the fast optimization of multi-gene pathways with different gene configurations inside a specific host (Jones et al., 2015). One of the requirements of synthetic assembly is the orthogonality, i.e., the lack of interferences coming from biochemical interactions between artificial elements (Michener et al., 2012). In this context, fluorescent reporter system can be used for the fine-tuning of the different gene expression systems required for the artificial pathway (Schenzielorz, 2013; Nikel et al., 2014a). This helps finding the best vector configurations allowing the balanced expression of all the enzymes involved in the artificial pathway.



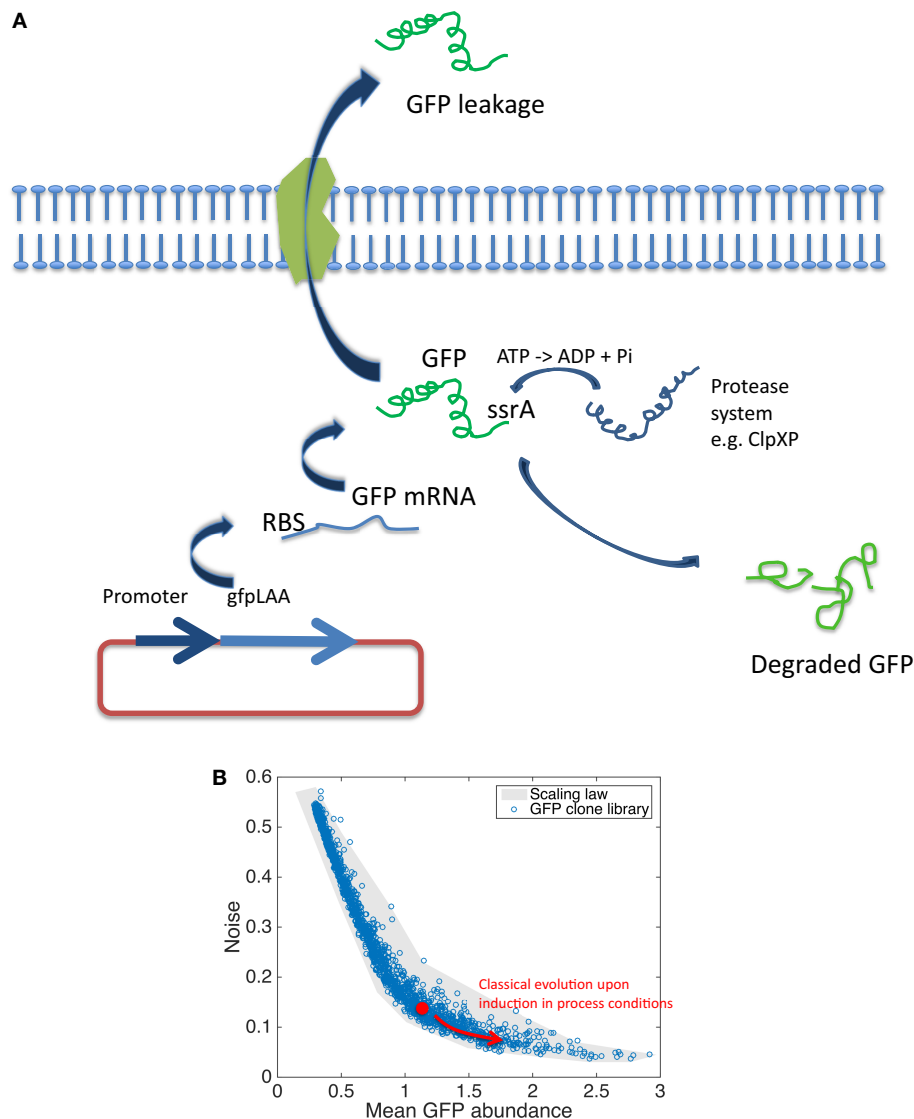
## Design of Complex Phenotypes for Food and White Biotechnology Applications: Directed and Laboratory Evolution

All of the above-mentioned examples imply the use of genetically modified strains. However, there are a lot of biotechnological applications where GMOs are not allowed, such as for agro-food applications for the development of starters and probiotics. More generally, there are a lot of biotechnological applications where complex phenotypes, combining robustness and productivity of the microbial cell factory, are needed. Such phenotypes are needed for the development of white biotechnology applications, i.e., the production of fuels and chemicals from complex lignocellulosic substrates containing inhibitors (Vasdekis and Stephanopoulos, 2015). In this context, the fluorescent reporter technology would give also invaluable insight about the phenotypic differentiation leading to the appearance of such phenotypes. Indeed, fitness is linked with noise in biochemical processes, and noise can be studied by the use of specific fluorescent reporter (Abee et al., 2011; Ryall et al., 2012; Holland et al., 2013). These reporters can be used to study bet hedging, a phenomenon by which a microbial population exploit phenotypic noise in order to increase its fitness in an ecosystem (Veening, 2008). Bet hedging has been recently identified as a major mechanism in diauxic shift (Boulineau et al., 2013; Solopova et al., 2014; van Heerden, 2014). Since diauxic shift arises frequently in bioprocesses, and especially within those based on complex substrate, it is of importance to control this mechanism. On the other hand, the design of complex phenotypes can be obtained by directed or natural evolution of a microbial strain (Dragosits and Mattanovich, 2013). At this level, also fluorescent reporters can be used to visualize in real time laboratory evolution (VERT). The VERT system is based on competitive fitness principle with different strains placed simultaneously in the cultivation system (Reyes, 2012a,b). The best strain can be easily selected since each strain carries a fluorescent reporter exhibiting a specific color. This technique allows for the fast detection of evolved phenotypes, the reporter plasmid being cured at the end of the experiments in order to meet the non-GMO requirements.

## Drawbacks Associated with Promoter-Based Fluorescent Biosensors

Reporter-based fluorescent biosensors have been widely used in the context of bioprocess optimization (Polizzi and Kontoravdi, 2014). However, some precautions must be taken since several factors can have a significant impact on the synthesis of the reporter molecules. These factors are summarized at **Figure 2A** and comprises the plasmid copy number, the strength of ribosome binding site (RBS), the folding rate of the GFP and its stability, and the potential release of GFP to the extracellular medium. Plasmid copy number is known to affect the degree of expression of GFP since this number can vary during cultures according to many environmental and intrinsic conditions. However, it has been shown that this effect can be limited by using low-copy number plasmids or by considering chromosomal integration (Freed et al., 2008; Silander et al., 2012). The strength of the RBS can also affect the rate by which mRNA is translated into GFP and this factor is

now integrated when designing synthetic gene circuits (Xu et al., 2012; Jones et al., 2015). Typical folding rate of engineered GFP (e.g., GFPmut2 and GFPmut3) are below 4 min (Cormack, 1996), but the stability is very high (>24 h) leading to cumulative signal. Destabilized version of GFP have been designed and used to monitor gene expression in bioprocessing conditions (Han et al., 2013; Hentschel et al., 2013). Destabilization of GFP relies on the use of specific *ssrA* tags that can be recognized by the internal protease machinery (i.e., *ClpXP* mainly). By varying the sequence of the last three amino acids, the affinity of the proteases can be modulated, leading to medium (e.g., GFPAAV for *E. coli* with a half-life of 40 min) to highly destabilized variants (e.g., GFPLAA for *E. coli* with a half-life of 10 min). However, in this case, the GFP response is affected not only by the real promoter activity but also by the content of intracellular protease and the adenosine triphosphate (ATP) availability since *ClpXP* is ATP dependent (Purcell et al., 2012; Han et al., 2013). These side-reactions can notably artificially enhance the cell-to-cell variability at the level of the GFP content (Baert et al., 2015). A last effect potentially affecting the overall level of GFP is its potential release to the extracellular medium. Indeed, increase in membrane permeability and protein leakage is known to occur when microbial cells are exposed to nutrient limitation, a condition encountered in fed-batch bioprocesses (Shokri, 2002, 2004). At this level, it has been shown that GFP can be subjected to such release (Delvigne et al., 2011; Brognaux et al., 2014). In the case of promoter-dependent reporter systems, all these side-reactions affect the reliability of the fluorescent signal according to the real activity of the targeted promoter. Also, the impact of cultivation conditions can also have a significant impact both on the mean expression level of GFP and on its cell-to-cell variability (called noise). Noise in GFP expression is indeed a very important factor that can have detrimental effect on productivity (Delvigne and Goffin, 2014; Delvigne et al., 2014). However, noise in gene expression is also recognized as a beneficial factor in the context of microbial robustness. Indeed, cell-to-cell heterogeneity increases the potential resistance of the microbial population upon stress exposure by a mechanism named bet-hedging (Holland et al., 2013; Nikel et al., 2014b; Martins and Locke, 2015). Since productivity and robustness are two traits that are generally expected from an efficient cell factory, the effect of noise have to be balanced between these two phenotypic traits and advanced analytical devices are awaited at this level (see **Figure 1** for a description of such devices). At this level, the use of an *E. coli* GFP clone library (Zaslaver et al., 2006) in connection with high-throughput flow cytometry has allowed for the design of a scaling law (**Figure 2B**). The collection comprised more than 1500 promoter-based biosensors and has been used for a genome-wide screen of the relationship between the promoter activity and the noise (Silander et al., 2012). The scaling-law shows a clear relationship between the mean promoter activity and the noise in GFP expression, i.e., promoters with high expression level exhibiting a low amount of noise, whereas weak promoters tend to exhibit larger level of noise (**Figure 2B**). More recently, this scaling-law has been validated by on-line flow cytometry in real bioprocessing conditions (Baert et al., 2015). This scaling-law constitutes an invaluable tool for a better understanding of the expression efficiency of microbial systems.



**FIGURE 2 | (A)** Scheme showing the different molecular mechanisms affecting the intracellular content in GFP. **(B)** Scaling law adapted from Silander et al. (2012) showing the relationship between the level of promoter activity (in log scale) and its cell-to-cell variability or noise (the noise is calculated as the ratio between the standard deviation and the mean GFP content among a population of microbial cell, in log scale). The red arrow shows a typical evolution in GFP expression in real bioprocessing conditions (Baert et al., 2015).

The study of microbial physiology based on the use of fluorescent reporter system is thus critical in order to increase our knowledge for the development of more efficient microbial cell factories. However, this development is slowed down by the development of cultivation devices compatible with single-cell analysis and by the need for reliable fluorescent reporter systems (Polizzi and Kontoravdi, 2014). Indeed, we have shown that promoter-based fluorescent reporters are subject to many side-reactions affecting the interpretation of the fluorescent signal. Efforts must thus be given at the level of the development of more reliable expression systems. Promoter-independent fluorescent reporter systems described in the Section “Detection of Metabolite Productivity at the Intracellular Level” and the developments of synthetic devices

described in the Section “Synthetic Biology and the Optimization of Artificial Metabolic Pathways” are very promising at this level.

## Promises of Promoter-Independent Fluorescent Biosensors

According to previously established classifications, three types of fluorescent biosensors can be distinguished, i.e., promoter-based, transcriptional regulator, and RNA switches (Eggeling et al., 2015; Zhang et al., 2015). Besides these three main classes, a fourth one can also be considered, i.e., fluorescent protein based biosensors (Delvigne and Goffin, 2014; Liu et al., 2015). This type of biosensor relies on the constitutive expression of fluorescent protein linked

with metabolite-binding domain and can be used for the intracellular detection of small molecules, such as ATP, cyclic adenosine mono-phosphate (cAMP), sugars, and organic acids, among others. Adequate constitutive expression of the fluorescent protein allows to get rid of most of the drawbacks depicted at **Figure 2A** and leads to a very dynamics signal that can be easily recorded by standard techniques (**Figure 1**). Additional advantages of this simple technology are the direct detection of important intracellular metabolites and its applicability to a broad range of hosts (Delvigne and Goffin, 2014). RNA switches, or riboswitches, can also be applied in the same context. These biosensors are based on mRNA elements able to sense metabolite concentration and modulate transcription and translation of a fluorescent protein accordingly. Their main drawback is the limited number of applications available so far (Eggeling et al., 2015), partially attributed to the fact that only a few number of natural mRNA regions with specific ligand properties are available. However, synthetic biology approaches allow for the design of artificial riboswitches with ligand capabilities for virtually all possible metabolites to be detected (Pothoulakis et al., 2014; Jang et al., 2015; You et al., 2015). Transcription-based fluorescent biosensors are based on the natural capabilities for transcription factor to sense protein interaction or change in intracellular metabolite concentration and regulate gene expression accordingly. This property can be used in order to drive the expression of a fluorescent protein in front of specific stimuli. This class of biosensor has been thoroughly used for monitoring the concentration of intracellular compounds (i.e., metabolites, co-factors, etc.) in metabolic engineering studies. However, since it relies on the synthesis of fluorescent proteins, it suffers from the same molecular drawbacks than classical promoter-based biosensors (**Figure 2A**). Another drawback is its host specificity. However, synthetic biology has been recently used for the design of more robust and widely applicable transcriptional biosensors (Zhang et al., 2015).

## References

- Abee, T., Wels, M., de Been, M., and den Besten, H. (2011). From transcriptional landscapes to the identification of biomarkers for robustness. *Microb. Cell Fact.* 10(Suppl. 1), S9. doi:10.1186/1475-2859-10-S1-S9
- Abu-Absi, N. R., Zamamiri, A., Kacmar, J., Balogh, S. J., and Srien, F. (2003). Automated flow cytometry for acquisition of time-dependent population data. *Cytometry A* 51, 87–96. doi:10.1002/cyto.a.10016
- Arnoldini, M., Heck, T., Blanco-Fernandez, A., and Hammes, F. (2013). Monitoring of dynamic microbiological processes using real-time flow cytometry. *PLoS ONE* 8:e80117. doi:10.1371/journal.pone.0080117
- Baert, J., Kinet, R., Brognaux, A., Delepierre, A., Telek, S., Sorensen, S. J., et al. (2015). Phenotypic variability in bioprocessing conditions can be tracked on the basis of on-line flow cytometry and fits to a scaling law. *Biotechnol. J.* 10, 1316–1325. doi:10.1002/biot.201400537
- Besmer, M. D., Weissbrodt, D. G., Kratochvil, B. E., Sigrist, J. A., Weyland, M. S., and Hammes, F. (2014). The feasibility of automated online flow cytometry for in-situ monitoring of microbial dynamics in aquatic ecosystems. *Front. Microbiol.* 5:265. doi:10.3389/fmicb.2014.00265
- Binder, S. (2012). A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. *Genome Biol.* 13, R40. doi:10.1186/gb-2012-13-5-r40
- Binder, S. (2013). Recombineering in *Corynebacterium glutamicum* combined with optical nanosensors: a general strategy for fast producer strain generation. *Nucleic Acids Res.* 41, 6360–6369. doi:10.1093/nar/gkt312

## Conclusion and Future Perspectives

The different examples shown in the previous sections point out that fluorescent reporter libraries have become a useful experimental tool used in different contexts for the optimization of the microbial cell factories, not only in the field of recombinant products but also for the production of metabolites, i.e., for detection of highly productive phenotypes and the orientation of the synthetic biology approaches.

In addition, the panel of experimental devices available for the analysis of such library is very large (**Figure 1**) and appropriate combination of such device on the basis of fluorescent reporter libraries allows for the high-throughput development of cell factories, from the selection of the best producers to the optimization of large-scale processes. However, additional efforts must be given in order to better integrate phenotypic noise. Indeed, cell-to-cell variation is an important factor known to affect bioprocess productivity and a specific mathematical background have to be set up in order to characterize this important phenomenon.

We have shown that most of the studies published actually are based on promoter-dependent fluorescent reporters. These molecular devices exhibit a lot of artifact that can lead to strong variation in fluorescent molecules expression. At this level, the development of promoter-independent fluorescent reporter would lead to more reliable clone libraries that can be used for the fast optimization of microbial cell factories.

## Acknowledgments

This project has received European Regional Development Funding through INTERREG IVB “Investing in opportunities.” This work was supported by the BioRefine Project (INTERREG IVB NWE Programme) (ref. 320J-BIOREFINE) and the RENEW Project (ref. 317J-RENEW).

- Boulineau, S., Tostevin, F., Kiviet, D. J., ten Wolde, P. R., Nghe, P., and Tans, S. J. (2013). Single-cell dynamics reveals sustained growth during diauxic shifts. *PLoS ONE* 8:e61686. doi:10.1371/journal.pone.0061686
- Brognaux, A., Francis, F., Twizere, J. C., Thonart, P., and Delvigne, F. (2014). Scale-down effect on the extracellular proteome of *Escherichia coli*: correlation with membrane permeability and modulation according to substrate heterogeneities. *Bioprocess Biosyst. Eng.* 14, 14. doi:10.1007/s00449-013-1119-8
- Brognaux, A., Han, S., Sorensen, S. J., Lebeau, F., Thonart, P., and Delvigne, F. (2013). A low-cost, multiplexable, automated flow cytometry procedure for the characterization of microbial stress dynamics in bioreactors. *Microb. Cell Fact.* 12, 100. doi:10.1186/1475-2859-12-100
- Cabantous, S., Terwilliger, T. C., and Waldo, G. S. (2005). Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat. Biotechnol.* 23, 102–107. doi:10.1038/nbt1044
- Cormack, B. P. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* 173, 33–38. doi:10.1016/0378-1119(95)00685-0
- DeLisa, M. P. (1999). Monitoring GFP operon fusion protein expression during high cell density cultivation of *Escherichia coli* using an on-line optical sensor. *Biotechnol. Bioeng.* 65, 54–64. doi:10.1002/(SICI)1097-0290(19991005)65:1<54::AID-BIT7>3.3.CO;2-I
- Delvigne, F., Baert, J., Gofflot, S., Lejeune, A., Telek, S., Johanson, T., et al. (2015). Dynamic single-cell analysis of *Saccharomyces cerevisiae* under process perturbation: comparison of different methods for monitoring the intensity of population heterogeneity. *J. Chem. Technol. Biotechnol.* 90, 314–323. doi:10.1002/jctb.4430

- Delvigne, F., Brognaux, A., Francis, F., Twizere, J. C., Gorret, N., Sorensen, S. J., et al. (2011). Green fluorescent protein (GFP) leakage from microbial biosensors provides useful information for the evaluation of the scale-down effect. *Biotechnol. J.* 6, 968–978. doi:10.1002/biot.201000410
- Delvigne, F., and Goffin, P. (2014). Microbial heterogeneity affects bioprocess robustness: dynamic single cell analysis contribute to understanding microbial populations. *Biotechnol. J.* 9, 61–72. doi:10.1002/biot.201300119
- Delvigne, F., Zune, Q., Lara, A. R., Al-Soud, W., and Sorensen, S. J. (2014). Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. *Trends Biotechnol.* 32, 608–616. doi:10.1016/j.tibtech.2014.10.002
- Dragosits, M., and Mattanovich, D. (2013). Adaptive laboratory evolution – principles and applications for biotechnology. *Microb. Cell Fact.* 12, 64. doi:10.1186/1475-2859-12-64
- Eggeling, L., Bott, M., and Marienhagen, J. (2015). Novel screening methods-biosensors. *Curr. Opin. Biotechnol.* 35C, 30–36. doi:10.1016/j.copbio.2014.12.021
- Freed, N. E., Silander, O. K., Stecher, B., Bohm, A., Hardt, W.-D., and Ackermann, M. (2008). A simple screen to identify promoters conferring high levels of phenotypic noise. *PLoS Genet.* 4:e1000307. doi:10.1371/journal.pgen.1000307
- Funke, M., Buchenauer, A., Mokwa, W., Kluge, S., Hein, L., Muller, C., et al. (2010). Bioprocess control in microscale: scalable fermentations in disposable and user-friendly microfluidic systems. *Microb. Cell Fact.* 9, 86. doi:10.1186/1475-2859-9-86
- Grunberger, A., Paczia, N., Probst, C., Schendzielorz, G., Eggeling, L., Noack, S., et al. (2012). A disposable picolitre bioreactor for cultivation and investigation of industrially relevant bacteria on the single cell level. *Lab. Chip* 12, 2060–2068. doi:10.1039/c2lc40156h
- Grunberger, A., van Ooyen, J., Paczia, N., Rohe, P., Schindzielorz, G., Eggeling, L., et al. (2013). Beyond growth rate 0.6: *Corynebacterium glutamicum* cultivated in highly diluted environments. *Biotechnol. Bioeng.* 110, 220–228. doi:10.1002/bit.24616
- Grunberger, A., Wiechert, W., and Kohlheyer, D. (2014). Single-cell microfluidics: opportunity for bioprocess development. *Curr. Opin. Biotechnol.* 29, 15–23. doi:10.1016/j.copbio.2014.02.008
- Hai, M., and Magdassi, S. (2004). Investigation on the release of fluorescent markers from w/o/w emulsions by fluorescence-activated cell sorter. *J. Control Release* 96, 393–402. doi:10.1016/j.jconrel.2004.02.014
- Han, S., Delvigne, F., Brognaux, A., Charbon, G. E., and Sorensen, S. J. (2013). Design of growth-dependent biosensors based on destabilized GFP for the detection of physiological behavior of *Escherichia coli* in heterogeneous bioreactors. *Biotechnol. Prog.* 29, 553–563. doi:10.1002/btpr.1694
- Hentschel, E., Will, C., Mustafi, N., Burkovski, A., Rehm, N., and Frunzke, J. (2013). Destabilized eYFP variants for dynamic gene expression studies in *Corynebacterium glutamicum*. *Microb. Biotechnol.* 6, 196–201. doi:10.1111/j.1751-7915.2012.00360.x
- Holland, S. L., Reader, T., Dyer, P. S., and Avery, S. V. (2013). Phenotypic heterogeneity is a selected trait in natural yeast populations subject to environmental stress. *Environ. Microbiol.* 11, 1729–1740. doi:10.1111/1462-2920.12243
- Jang, S., Yang, J., Seo, S. W., and Jung, G. Y. (2015). Riboselector: riboswitch-based synthetic selection device to expedite evolution of metabolite-producing microorganisms. *Meth. Enzymol.* 550, 341–362. doi:10.1016/bs.mie.2014.10.039
- Jones, J. A., Vernacchio, V. R., Lachance, D. M., Lebovich, M., Fu, L., Shirke, A. N., et al. (2015). ePathOptimize: a combinatorial approach for transcriptional balancing of metabolic pathways. *Sci. Rep.* 5, 11301. doi:10.1038/srep11301
- Kintses, B., Hein, C., Mohamed, M. F., Fischlechner, M., Courtois, F., Laine, C., et al. (2012). Picoliter cell lysate assays in microfluidic droplet compartments for directed enzyme evolution. *Chem. Biol.* 19, 1001–1009. doi:10.1016/j.chembiol.2012.06.009
- Klockner, W., and Buchs, J. (2012). Advances in shaking technologies. *Trends Biotechnol.* 30, 307–314. doi:10.1016/j.tibtech.2012.03.001
- Lattermann, C., and Buchs, J. (2014). Microscale and miniscale fermentation and screening. *Curr. Opin. Biotechnol.* 35C, 1–6. doi:10.1016/j.copbio.2014.12.005
- Liu, D., Evans, T., and Zhang, F. (2015). Applications and advances of metabolite biosensors for metabolic engineering. *Metab. Eng.* 31, 35–43. doi:10.1016/j.ymben.2015.06.008
- Love, K. R. (2010). Integrated single-cell analysis shows *Pichia pastoris* secretes protein stochastically. *Biotechnol. Bioeng.* 106, 319–325. doi:10.1002/bit.22688
- Love, K. R. (2012). Systematic single-cell analysis of *Pichia pastoris* reveals secretory capacity limits productivity. *PLoS ONE* 7:e37915. doi:10.1371/journal.pone.0037915
- Love, K. R. (2013). Microtools for single-cell analysis in biopharmaceutical development and manufacturing. *Trends Biotechnol.* 31, 280–286. doi:10.1016/j.tibtech.2013.03.001
- Martinez-Garcia, E., Nikel, P. I., Aparicio, T., and de Lorenzo, V. (2014). *Pseudomonas* 2.0: genetic upgrading of *P. putida* KT2440 as an enhanced host for heterologous gene expression. *Microb. Cell Fact.* 13, 159. doi:10.1186/s12934-014-0159-3
- Martins, B. M., and Locke, J. C. (2015). Microbial individuality: how single-cell heterogeneity enables population level strategies. *Curr. Opin. Microbiol.* 24, 104–112. doi:10.1016/j.mib.2015.01.003
- Mazutis, L., Gilbert, J., Ung, W. L., Weitz, D. A., Griffiths, A. D., and Heyman, J. A. (2013). Single-cell analysis and sorting using droplet-based microfluidics. *Nat. Protoc.* 8, 870–891. doi:10.1038/nprot.2013.046
- Michener, J. K., Thodey, K., Liang, J. C., and Smolke, C. D. (2012). Applications of genetically-encoded biosensors for the construction and control of biosynthetic pathways. *Metab. Eng.* 14, 212–222. doi:10.1016/j.ymben.2011.09.004
- Newman, J. R. S., Ghaemmaghami, S., Ihmels, J., Breslow, D. K., Noble, M., DeRisi, J. L., et al. (2006). Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* 441, 840–846. doi:10.1038/nature04785
- Nikel, P. I., Martinez-Garcia, E., and de Lorenzo, V. (2014a). Biotechnological domestication of pseudomonads using synthetic biology. *Nat. Rev. Microbiol.* 12, 368–379. doi:10.1038/nrmicro3253
- Nikel, P. I., Silva-Rocha, R., Benedetti, I., and de Lorenzo, V. (2014b). The private life of environmental bacteria: pollutant biodegradation at the single cell level. *Environ. Microbiol.* 16, 628–642. doi:10.1111/1462-2920.12360
- Polizzi, K. M., and Kontoravdi, C. (2014). Genetically-encoded biosensors for monitoring cellular stress in bioprocessing. *Curr. Opin. Biotechnol.* 31, 50–56. doi:10.1016/j.copbio.2014.07.011
- Pothoulakis, G., Ceroni, F., Reeve, B., and Ellis, T. (2014). The spinach RNA aptamer as a characterization tool for synthetic biology. *ACS Synth. Biol.* 3, 182–187. doi:10.1021/sb400089c
- Purcell, O., Grierson, C. S., Bernardo, M. D., and Savery, N. J. (2012). Temperature dependence of ssrA-tag mediated protein degradation. *J. Biol. Eng.* 6, 10. doi:10.1186/1754-1611-6-10
- Reyes, L. H. (2012a). Visualizing evolution in real time to determine the molecular mechanisms of n-butanol tolerance in *Escherichia coli*. *Metab. Eng.* 14, 579–590. doi:10.1016/j.ymben.2012.05.002
- Reyes, L. H. (2012b). Visualizing evolution in real-time method for strain engineering. *Front. Microbiol.* 3:198. doi:10.3389/fmicb.2012.00198
- Rokke, G., Korvald, E., Pahr, J., Oyas, O., and Lale, R. (2014). BioBrick assembly standards and techniques and associated software tools. *Methods Mol. Biol.* 1116, 1–24. doi:10.1007/978-1-62703-764-8\_1
- Ryall, B., Eyddallin, G., and Ferenci, T. (2012). Culture history and population heterogeneity as determinants of bacterial adaptation: the adaptomics of a single environmental transition. *Microbiol. Mol. Biol. Rev.* 76, 597–625. doi:10.1128/MMBR.05028-11
- Sanchez, A., Choubey, S., and Kondev, J. (2013). Regulation of noise in gene expression. *Annu. Rev. Biophys.* 42, 469–491. doi:10.1146/annurev-biophys-083012-130401
- Schallmey, M., Frunzke, J., Eggeling, L., and Marienhagen, J. (2014). Looking for the pick of the bunch: high-throughput screening of producing microorganisms with biosensors. *Curr. Opin. Biotechnol.* 26, 148–154. doi:10.1016/j.copbio.2014.01.005
- Schendzielorz, G. (2013). Taking control over control: use of product sensing in single cells to remove flux control at key enzymes in biosynthesis pathways. *ACS Synth. Biol.* 15, 15. doi:10.1021/sb400059y
- Shokri, A. (2002). Growth rate dependent changes in *Escherichia coli* membrane structure and protein leakage. *Appl. Microbiol. Biotechnol.* 58, 386–392. doi:10.1007/s00253-001-0889-0
- Shokri, A. (2004). Characterisation of the *Escherichia coli* membrane structure and function during fedbatch cultivation. *Microb. Cell Fact.* 3, 9. doi:10.1186/1475-2859-3-9
- Silander, O. K., Nikolic, N., Zaslaver, A., Bren, A., Kikoin, I., Alon, U., et al. (2012). A genome-wide analysis of promoter-mediated phenotypic noise in *Escherichia coli*. *PLoS Genet.* 8:e1002443. doi:10.1371/journal.pgen.1002443



- Silva-Rocha, R., and de Lorenzo, V. (2010). Noise and robustness in prokaryotic regulatory networks. *Annu. Rev. Microbiol.* 64, 257–275. doi:10.1146/annurev.micro.091208.073229
- Solopova, A., van Gestel, J., Weissing, F. J., Bachmann, H., Teusink, B., Kok, J., et al. (2014). Bet-hedging during bacterial diauxic shift. *Proc. Natl. Acad. Sci. U.S.A.* 111, 7427–7432. doi:10.1073/pnas.1320063111
- Swain, P. S., Elowitz, M. B., and Siggia, E. D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 99, 12795–12800. doi:10.1073/pnas.162041399
- Taniguchi, Y., Choi, P. J., Li, G.-W., Chen, H., Babu, M., Hearn, J., et al. (2010). Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* 329, 533–538. doi:10.1126/science.1188308
- van Heerden, J. H. (2014). Lost in transition: startup of glycolysis yields subpopulations of nongrowing cells. *Science* 343, 1–9. doi:10.1126/science.1245114
- Vasdekis, A. E., and Stephanopoulos, G. (2015). Review of methods to probe single cell metabolism and bioenergetics. *Metab. Eng.* 27, 115–135. doi:10.1016/j.ymben.2014.09.007
- Veening, J. W. (2008). Bistability, epigenetics, and bet-hedging in bacteria. *Annu. Rev. Microbiol.* 62, 193–210. doi:10.1146/annurev.micro.62.081307.163002
- Xu, P., Vansiri, A., Bhan, N., and Koffas, M. A. G. (2012). ePathBrick: a synthetic biology platform for engineering metabolic pathways in *E. coli*. *ACS Synth. Biol.* 1, 256–266. doi:10.1021/sb300016b
- You, M., Litke, J. L., and Jaffrey, S. R. (2015). Imaging metabolite dynamics in living cells using a Spinach-based riboswitch. *Proc. Natl. Acad. Sci. U.S.A.* 112, E2756–E2765. doi:10.1073/pnas.1504354112
- Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., et al. (2006). A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nat. Methods* 3, 623–628. doi:10.1038/nmeth895
- Zaslaver, A., Mayo, A. E., Rosenberg, R., Bashkin, P., Sberro, H., Tsalyuk, M., et al. (2004). Just-in-time transcription program in metabolic pathways. *Nat. Genet.* 36, 486–491. doi:10.1038/ng1348
- Zhang, J., Jensen, M. K., and Keasling, J. D. (2015). Development of biosensors and their application in metabolic engineering. *Curr. Opin. Chem. Biol.* 28, 1–8. doi:10.1016/j.cbpa.2015.05.013

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Delvigne, Pêcheux and Tarayre. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Modifying Yeast Tolerance to Inhibitory Conditions of Ethanol Production Processes

Luis Caspeta<sup>1\*</sup>, Tania Castillo<sup>1</sup> and Jens Nielsen<sup>2,3,4</sup>

<sup>1</sup> Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos, Cuernavaca, Mexico,

<sup>2</sup> Novo Nordisk Foundation Center for Biosustainability, Chalmers University of Technology, Gothenburg, Sweden,

<sup>3</sup> Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden, <sup>4</sup> Novo Nordisk Foundation Center for Biosustainability, Hørsholm, Denmark

## OPEN ACCESS

### Edited by:

Hilal Taymaz Nikerel,  
Bogazici University, Turkey

### Reviewed by:

Caroline Evans,  
University of Sheffield, UK  
Steven W. Gorsich,  
Central Michigan University, USA  
Nianshu Zhang,  
University of Cambridge, UK

### \*Correspondence:

Luis Caspeta  
luis.caspeta@uaem.mx

### Specialty section:

This article was submitted to  
Systems Biology,  
a section of the journal  
Frontiers in Bioengineering and  
Biotechnology

**Received:** 16 August 2015

**Accepted:** 28 October 2015

**Published:** 11 November 2015

### Citation:

Caspeta L, Castillo T and Nielsen J  
(2015) Modifying Yeast Tolerance to  
Inhibitory Conditions of Ethanol  
Production Processes.  
Front. Bioeng. Biotechnol. 3:184.  
doi: 10.3389/fbioe.2015.00184

*Saccharomyces cerevisiae* strains having a broad range of substrate utilization, rapid substrate consumption, and conversion to ethanol, as well as good tolerance to inhibitory conditions are ideal for cost-competitive ethanol production from lignocellulose. A major drawback to directly design *S. cerevisiae* tolerance to inhibitory conditions of lignocellulosic ethanol production processes is the lack of knowledge about basic aspects of its cellular signaling network in response to stress. Here, we highlight the inhibitory conditions found in ethanol production processes, the targeted cellular functions, the key contributions of integrated -omics analysis to reveal cellular stress responses according to these inhibitors, and current status on design-based engineering of tolerant and efficient *S. cerevisiae* strains for ethanol production from lignocellulose.

**Keywords:** yeast, stress tolerance, cellular stress response, inhibitory environment, ethanol production process, design-based engineering, integrated -omics analysis

## INTRODUCTION

Microbial fermentation of sugars from sugarcane and corn starch to ethanol is the source of around 100 billion liters of fuel ethanol annually produced in the world using *Saccharomyces cerevisiae*. There is also much interest in the use of lignocellulose as a feedstock for future production of ethanol, since this source is much more abundant and, most important, it does not compete with food for supplies (Lynd et al., 1991; Caspeta et al., 2013). However, one of the major bottlenecks for lignocellulose conversion to ethanol is that the related production processes should be economically competitive, a condition that is in detriment of yeast performance, since it must face high concentrations of toxic chemicals and harmful process conditions, for which extra operations for process conditioning to yeast tolerance are economically and energetically prohibited (Caspeta et al., 2013, 2014a). Thus, yeast cells can be exposed to inhibitory concentrations of toxic chemicals and low pH resulted from thermo-chemical pretreatment of lignocellulose. Furthermore, saccharification and fermentation of sugar polymers exposed *S. cerevisiae* to high temperatures, elevated osmolarity, and high concentrations of ethanol (Garay-Arroyo et al., 2004; Caspeta et al., 2014a). The former conditions are useful to reduce contamination and cooling efforts as well as to decrease energy utilization during downstream processing and to decrease enzyme loadings concomitant with lower production costs (Caspeta et al., 2014a).

Microorganisms capable of resisting conditions of lignocellulose ethanol production processes whereas maintaining high metabolic activity are desirable. Microbial strains with these characteristics

can be isolated from natural habitats where they have been evolving these traits for a long time (Ballesteros et al., 1991; Edgardo et al., 2008; Field et al., 2015). Another option is to generate tolerant phenotypes in model organisms like *S. cerevisiae*. This requires the augmentation of limits between the relation of cellular functions and environmental fluctuations, namely to diminish the disturbing effects of inhibitory conditions on cellular functions required for growing and biofuel synthesis. Some of the targeted functions include proteome structure and stability, RNA synthesis and processing, sugar transport, membrane fluidity, and DNA processing, among others (Kültz, 2005).

Yeast and other microorganisms have gene expression and metabolic turnover programs that have been finely adjusted to improve cells fitness in the environmental fluctuations found in their natural environments (Tagkopoulos et al., 2008; Mitchell et al., 2009). Therefore, microorganisms exposed to novel environments may mount erratic non-specific responses leading them to survive or perish. Thus, one can expect that adaptation to novel environments will require the complete reprogramming of cellular functions, including gene expression and metabolic turnover, which may not be attainable by multigene modification (Alper and Stephanopoulos, 2009). This is probably more evident due to the fact that stressful conditions – out of those found in the natural environments – will not be anticipated in native signaling networks. However, there is genomic plasticity that allows approaching the hypothesis that cells can acquire new functions or reconfigure macromolecular structures more suitable to new environments. The challenge then is to recognize the genomic rearrangements and the resultant levels of gene expression, according to environmental changes. In this review, we describe basic knowledge about cellular stress response (CSR) and the current strategies for improving yeast tolerance to inhibitory conditions found in lignocellulosic ethanol production processes.

## INHIBITORY CONDITIONS OF LIGNOCELLULOSIC ETHANOL PRODUCTION PROCESS

Lignocellulose is a tightly packed structure of the carbohydrate polymers cellulose and hemicellulose surrounded by the phenolic polymer lignin. Although several processes have been developed thus far for lignocellulose conversion to ethanol, a characteristic one includes the general steps shown in **Figure 1**. Once the material has been chopped in pieces, a pretreatment step mainly consisting of a thermo-chemical treatment of lignocellulose is used for its hydrolysis into fermentable sugars. These get dissolved in a syrup that can also contain acetic, formic, and levulinic acids, as well as furans and phenolic compounds released during pretreatment (Larsson et al., 1999; Palmqvist and Hahn-Hägerdal, 2000b). Since these chemicals reduce yeast growth and ethanol production (Zaldivar et al., 1999; Larsson et al., 2000; Palmqvist and Hahn-Hägerdal, 2000a), several efforts have been made to avoid their production (Palmqvist and Hahn-Hägerdal, 2000a; Caspeta et al., 2014a). Another option is to reduce their concentrations by different detoxification methods (Palmqvist and Hahn-Hägerdal,

2000a), but extra operations negatively impact energy balance and production costs (Caspeta and Nielsen, 2013).

Whatever the hydrolysis method, this must ensure syrups with high sugar concentrations. Concentrations of fermentable sugars higher than 250 g L<sup>-1</sup> guarantee ethanol titers above 100 g L<sup>-1</sup>, required to reduce energy consumption and production costs during downstream operations (Haelssig et al., 2008). To reach these concentrations, suspensions with around 416 g of pretreated lignocellulosic biomass containing 60% of fermentable sugars – a high gravity suspension will be needed. The resulted syrup would contain high amounts of toxic chemicals as well as elevated amounts of insoluble lignin and cellulose fractions. If saccharification and fermentation of cellulose is performed simultaneously, the high gravity of cellulose/lignin suspension could impair both, enzyme activity and cell growth (Caspeta et al., 2014a). Whereas, performing saccharification and fermentation separately exposes yeast cells to toxic compounds and very high osmolarity.

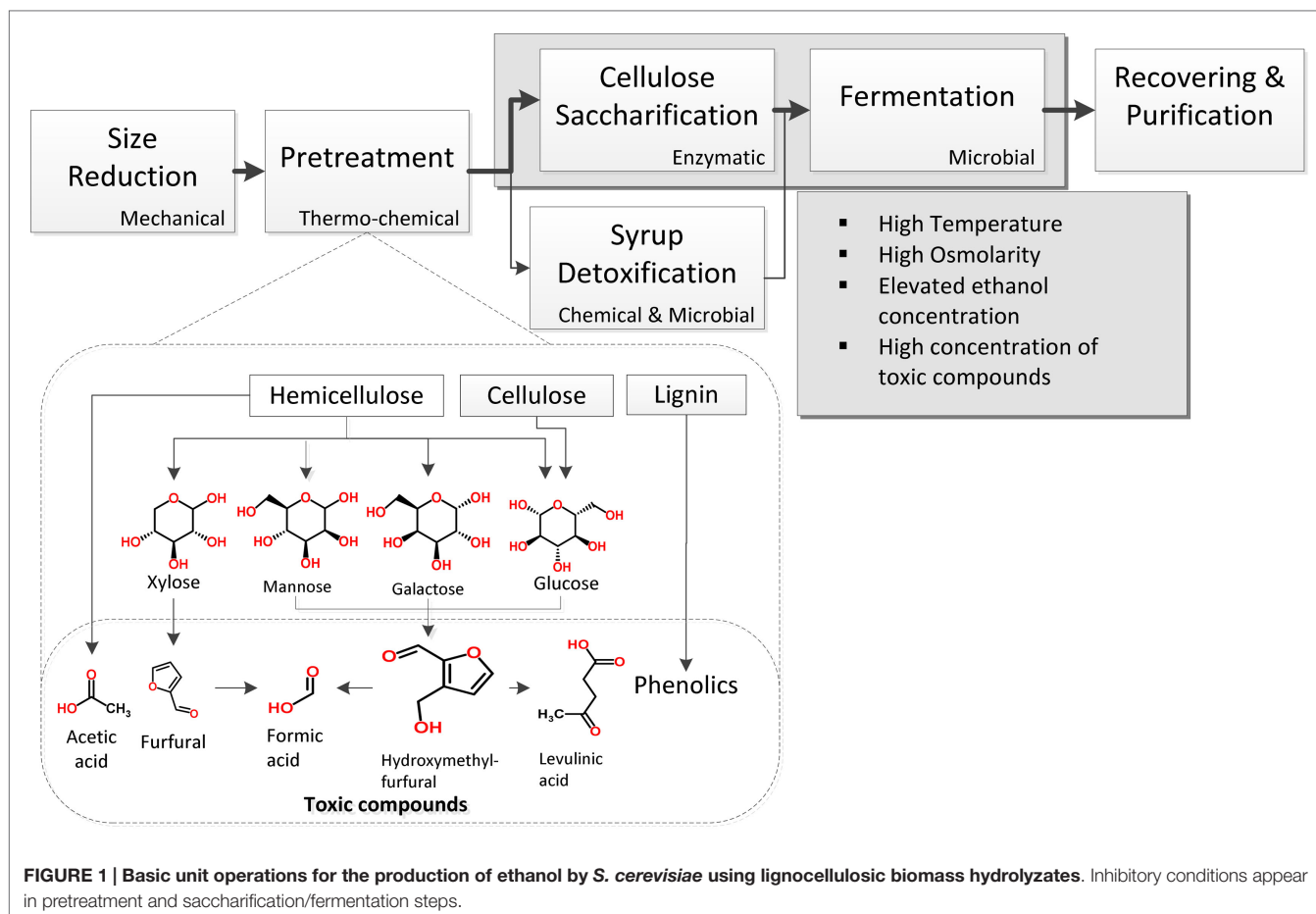
Performing thermo-chemical hydrolysis at mild conditions reduces toxic compounds formation and can disrupt lignocellulose structure (Pan et al., 2006; Caspeta et al., 2014a), keeping hemicellulose and/or cellulose polymers intact for their further hydrolysis with cellulosic enzymes. Saccharification is costly and highly affected by process temperature and solid loadings (Ingesson et al., 2001; Caspeta et al., 2014a). Most of commercial enzymes have optimal temperatures higher than 45°C and the enzymes' industry have been trying to increase it, because of operations at high temperatures are highly desirable to reduce contamination and cooling efforts. This condition, however, limits simultaneous saccharification and fermentation since most of yeast strains do not tolerate temperatures higher than 40°C.

In summary, *S. cerevisiae* can be exposed to a number of toxic compounds formed during pretreatment of biomass, e.g., low pH, unusual levels of sugar concentration and solid loadings in cellulose suspensions and hydrolyzates, lethal temperatures occurring in saccharification, and high ethanol concentrations resulting from the fermentation. All these inhibitory conditions affect cellular functions in the different forms as described below.

## INHIBITORY EFFECTS OF HARMFUL CONDITIONS OF LIGNOCELLULOSIC ETHANOL PRODUCTION PROCESS

### Inhibitory Effects of Toxic Compounds

The inhibition of cellular growth and metabolism by toxic compounds formed or released during hydrolysis of lignocellulosic biomass was detailed elsewhere (Palmqvist and Hahn-Hägerdal, 2000b), and summarized in **Table 1**. Harmfulness of acetic, formic, and levulinic acids depends on extracellular and intracellular pH, membrane permeability, and toxicity of the anionic forms of the acids (Palmqvist and Hahn-Hägerdal, 2000b; Maris et al., 2004). Once the acid goes into yeast cell, the intracellular pH drops and excessive proton accumulation is pumped out of the cells by various mechanisms, including proton translocation with the plasma membrane H<sup>+</sup>-ATPase mediated by ATP hydrolysis (Holyoak et al., 1996; Maris et al., 2004). This cellular process can



be very intensive in terms of ATP utilization. For example, in presence of sorbic, benzoic, and octanoic acids at pH 4.5, 5.0, and 4.0, respectively, a 10-, 4-, and 1.5-fold decrease in intracellular ATP levels can be observed due to increasing energy for maintenance of the internal pH (Viegas and Sá-Correia, 1991; Verduyn et al., 1992; Holyoak et al., 1996), with a concomitant reduction of biomass yields (Viegas and Sá-Correia, 1991; Verduyn et al., 1992). Furthermore, acetic and formic acids, in their anionic forms, are lipophobic and enter to the cell as undissociated forms, which prevail at external pH values below 4.8 (Casal et al., 1996). Inside the cell, the acid is dissociated and the intracellular pH decreases. It has been shown that intracellular concentrations higher than 120 mM of acetic acid reduce enolase and phosphoglyceromutase activities by 50% respect to non-acidic conditions (Pampulha and Loureiro-Dias, 1990). However, evidence suggests that proton exporting is the major contribution for reduced growth rate upon yeast exposition to acids.

The 5-hydroxymethyl furfural (HMF) and 2-furaldehyde (furfural) are formed from thermal oxidation of hexoses and pentoses during pretreatment (Palmqvist and Hahn-Hägerdal, 2000b) – **Figure 1**. These compounds induce chromatin changes, DNA damage, reduced translation, and inactivation of various glycolytic enzymes (Banerjee et al., 1981; Allen et al., 2010; Ask

et al., 2013a) (**Table 1**). Yeast can metabolize furfural and HMF to their less toxic alcohols by oxidoreductases using NAD(P)H as a cofactor, a metabolic process that occurs at high rates (Diaz De Villegas et al., 1992; Ask et al., 2013a). Their conversion increases the cellular energy for maintenance and reduces the concentration of redox cofactors (Taherzadeh et al., 1999; Sárvári Horváth et al., 2003; Ask et al., 2013a). Thus, this is associated to a reduction of glycerol production and an increase of acetate production during ethanol fermentation in the presence of furfural (Palmqvist et al., 1999a; Sárvári Horváth et al., 2003; Ask et al., 2013b). Results from exposing *S. cerevisiae* to these chemicals suggested that yeast growth is more sensitive to furfural than to HMF or high ethanol titers (Taherzadeh et al., 1999), because HMF has lower permeability and its conversion is less efficient than furfural (Larsson et al., 1999). Besides, accumulation of reactive oxygen species induced by furfural can damage *S. cerevisiae* mitochondrion and vacuole (Allen et al., 2010). Both compartments regulate redox balance of cytosol and losing their functions can result in a reduction of glucose consumption rates.

When mixtures of acetic acid and furfural are present in the fermentation, the specific growth rate decreased more than the sum of the individual effects (Palmqvist et al., 1999b), suggesting that cells expend higher amounts of energy for excreting



**TABLE 1 | Examples of negative effects of inhibitory conditions found in ethanol production processes on yeast performance.**

Stress	Negative effects in the yeast cells	Reference
Exposition to toxic compounds (furfural, HMF, and phenolic compounds)	Chromatin changes, DNA damage, and reduction of translation activity Enzyme inactivation Reduction of the intracellular concentrations of NAD(P)H Negative effects on sorting and signaling functions Reactive oxygen species formation Low biomass yields	Allen et al. (2010), Ask et al. (2013a)  Ask et al. (2013b) Keweloh et al. (1990) Larsson et al. (1999)
Exposition to organic acids	Reduction of biomass yields Decrease of the intracellular ATP levels, concomitant to an increase of the maintenance energy Drop of the intracellular pH Reduction of enzymatic activities	Viegas and Sá-Correia (1991)  Holoak et al. (1996) Pampulha and Loureiro-Dias (1990)
Exposition to ethanol	Impairment of cellular wall permeability Disruption of sorting and signaling functions, with an increment of the cell size Induction of petite mutants without mitochondrial DNA ( $\rho^0$ ) Reduction of metabolic activity Impairment of acid resistance	Kubota et al. (2004) Jones and Greenfield (1987)  Ibeas and Jimenez (1997) Nagodawithana and Steinkraus (1976) Pampulha and Loureiro-Dias (1989)
Osmotic	High accumulation of glycerol Accumulation of ethanol Disruption of actin cytoskeleton Disruption of MAP kinase cascade Reduction of cell viability	Hohmann (2002) D'Amore et al. (1988) Chowdhury et al. (1992)
Physicochemical (temperature and pH)	Augmentation of detrimental effects of toxic compounds Modification of the protein functional structure Reduction of enzymatic and metabolic activities Reduction of the cell growth	Piper (1993), Kültz (2005)

acid anions, protons, and furfural out of the cell, as well as for reactive oxygen species formed during furfural assimilation. The growth-inhibitory effects by potential lignocellulose-derived inhibitors, including phenols [lignin, vanillin, 4-hydroxybenzaldehyde (4-HB), and syringaldehyde], furans (furfural and 5-hydroxymethyl-2-furaldehyde), and organic acids (levulinic, formic, and acetic) on the growth and ethanol production were investigated. From these, phenols and furans exhibited potent inhibitory effects at a concentration of  $1 \text{ g L}^{-1}$ , while organic acids had insignificant impacts at concentrations of up to  $2 \text{ g L}^{-1}$ .

Phenolic compounds released from the hydrolysis of lignin are poorly soluble in aqueous solutions and they can be incorporated into cellular membranes where their partition is higher (Heipieper et al., 1994). Here, phenolic compounds mainly interfere with proteins function and trigger changes in the protein to lipid ratio (Keweloh et al., 1990). Hence, these compounds affect cellular functions like sorting and signaling, as well as cause membrane swelling. Among the 13 tested phenolic compounds, the 4-hydroxy-3-methoxycinnamaldehyde is the most toxic (Adeboye et al., 2014). This, vanillin and catechol are major constituents of syrups from pretreated lignocellulose (Ando, 1966; Palmqvist and Hahn-Hägerdal, 2000b). It is also abundant in hydrolyzates of hardwood, which is toxic at concentration of  $1 \text{ g L}^{-1}$ , reducing 30% of ethanol yield (Ando et al., 1986). The toxicity of phenolics is very variable as it depends on the functional groups (Ando et al., 1986; Jonsson et al., 2013; Adeboye et al., 2014); more methoxy groups are related to high hydrophobicity and toxicity (Klinke et al., 2004). *S. cerevisiae* can assimilate

many of phenolics which can be part of the detoxification process occurring during fermentation (Mills et al., 1971; Delgenes et al., 1996; Larsson et al., 2000).

## Inhibitory Effects of High Ethanol Concentrations

One of the main advantages of *S. cerevisiae* for ethanol production is the high tolerance that this yeast shows respect to other microorganisms. For example, whereas *Escherichia coli* and *Zymomonas mobilis* have maximum tolerances around  $60\text{--}127 \text{ g L}^{-1}$  (Lee et al., 1980; Yomano et al., 1998), *S. cerevisiae* can tolerate ethanol concentrations up to between 115 and  $200 \text{ g L}^{-1}$  (Luong, 1985). However, ethanol concentrations higher than  $150 \text{ g L}^{-1}$  can be required to reduce costs in downstream operations. High concentrations of alcohols like ethanol and butanol impaired cellular wall permeability disrupting sorting and signaling functions, as well as provoked an increase in cell size which caused a cell cycle delay (Jones and Greenfield, 1987; Kubota et al., 2004) (Table 1). This correlates with a dispersion of the F-actin cytoskeleton, which is probably regulated by the protein kinase SWE1, which regulates the G2/M transition, since its mutations abolish this phenotype (Kubota et al., 2004). Ethanol also induces petite mutants without mitochondrial DNA (the  $\rho^0$  mutants) and changes in mitochondrial genome (Ibeas and Jimenez, 1997; Chi and Arneborg, 1999). In combination with high temperature, ethanol exacerbates inactivation of some enzymes, for example, the alcohol dehydrogenase (ADH) and the

hexokinase (Augustin et al., 1965; Nagodawithana and Steinkraus, 1976; Chen and Jin, 2006). The uptake of alanine, proton efflux, and fermentation rates can decrease when cells are exposed to 2M of ethanol (Mishra and Prasad, 1989). Disruption of proton efflux also impairs acid resistance (Brown and Oliver, 1982; Sá-Correia and Van Uden, 1983; Gao and Fleet, 1988; Pampulha and Loureiro-Dias, 1989; Aguilera et al., 2006), since this affects proton outtake for regulation of internal pH. Interestingly, the activity of  $\beta$ -glucosidase, a cellulosic enzyme used in saccharification, increased with increasing ethanol concentrations from 1 to 9% (v/v) (Chen and Jin, 2006). Since cellular wall is the key ethanol target, yeast changes lipid composition, incrementing the proportion of polyunsaturated fatty acids (FAs), ergosterol, and phosphatidylcholine (Mishra and Prasad, 1989; Kajiwarra et al., 1996; Chi and Arneborg, 1999). This response is also observed in thermal stress. Eventually, moderate ethanol concentrations also reduce water activity with consequences in metabolic activity (Hallsworth, 1998).

### Inhibitory Effects of High Osmolarity

High gravity fermentations are required for economic considerations. Glucose concentrations superior to 300 g L<sup>-1</sup> are needed to reach ethanol titers higher than 150 g L<sup>-1</sup>. Thus, the osmolarity of a hydrolyzate can be of 20–200 g L<sup>-1</sup> of salt (0.6–8.6 Osm) (Olsson and Hahn-Hägerdal, 1993). *S. cerevisiae* can resist 4 Osm, which is much higher compared with *Z. mobilis*, which resists until 1.2 Osm. After being exposed to high osmolarity, yeast cells accumulated high amounts of glycerol which serves as an osmolyte (Hohmann, 2002). Osmotic shock disrupts actin cytoskeleton and invaginations appear affecting the conformation of actin bundles that disturbs MAP kinase cascade, which regulates cell cycle (Chowdhury et al., 1992). This also causes water to flow out of the cell, increasing the concentration of cellular components, including ion concentrations that can serve as a sensor for cellular signaling pathways (Hohmann, 2002). Under osmotic pressure, the excretion of ethanol and glycerol is impaired, leading the accumulation of intracellular ethanol and a decrease in cell viability (Panchal and Stewart, 1980; D'Amore et al., 1988). It seems that membrane fluidity is less prone to be affected by high osmolarity since medium pH does not have a significant effect on yeast growth at high glucose concentration, but only on ethanol accumulation (Narendranath and Power, 2005).

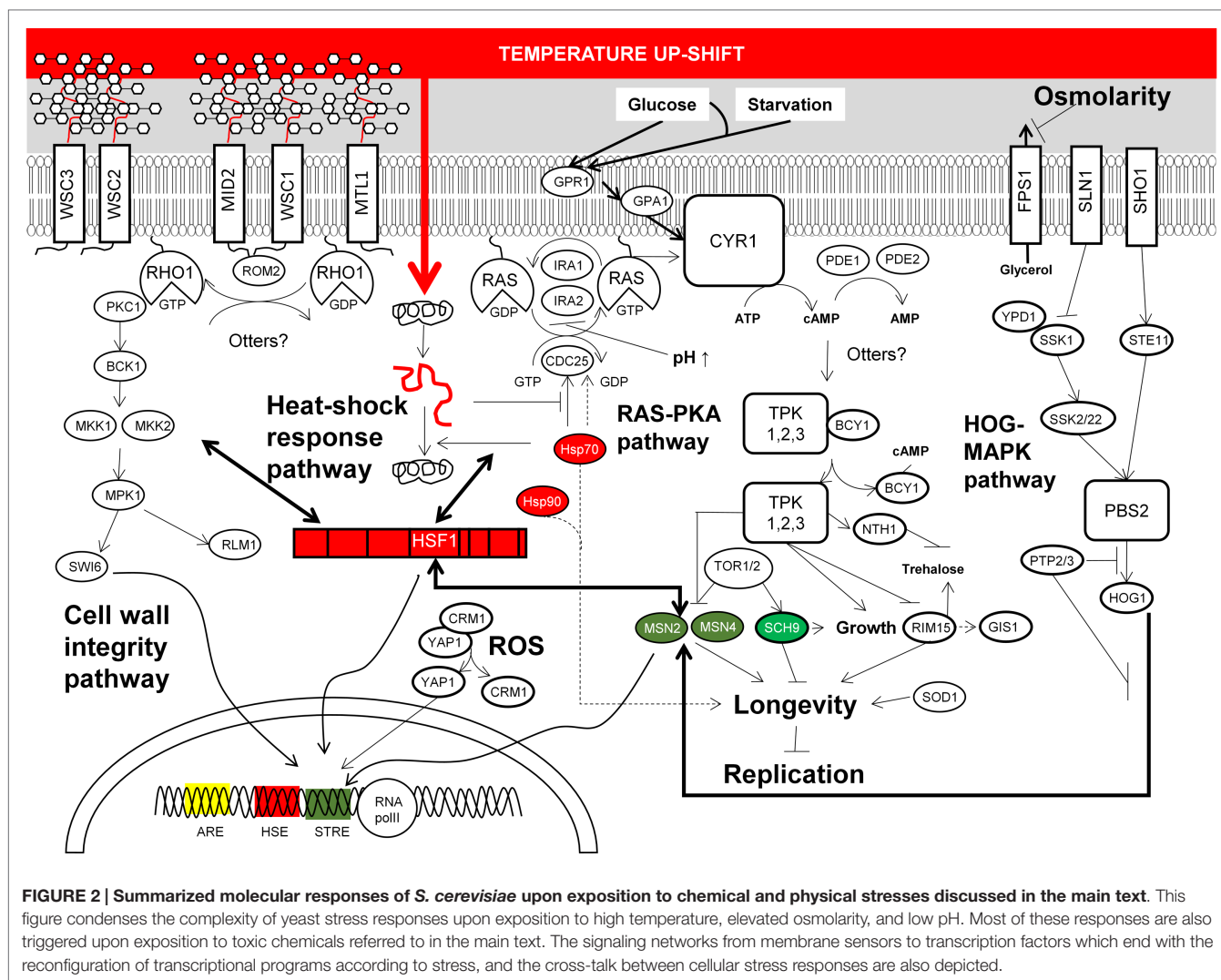
### Inhibitory Effects of High Temperature

Temperature pervasively practically affects all cellular macromolecules and metabolic functions (Table 1). Increasing temperature from 25–28°C to 40°C caused a substantial reduction of protein synthesis (Lindquist, 1981; Hottiger et al., 1987), which is accompanied by increasing trehalose accumulation (Hottiger et al., 1987; Neves and Francois, 1992). Both responses are essential to acquire thermotolerance (De Virgilio et al., 1994; Singer and Lindquist, 1998), since null mutants in the trehalose synthase (*TS1*) are more sensible to thermal stress (De Virgilio et al., 1994) and significantly decrease heat-shock genes transcription (Hazell et al., 1995), while cells carrying *CYR1-2* mutation produce trehalose constitutively, and are significantly more tolerant than the

wild type (Hottiger et al., 1989). There is also evidence that trehalose catabolism is needed to acquire thermotolerance and recovering of cellular homeostasis from thermal shock upon temperature upshift from 30 to 40°C (Nwaka et al., 1994, 1995). This is evidenced by the recovery of protein production and bud formation after starting trehalose degradation (Hottiger et al., 1987). Trehalose accumulates simultaneously with a reduction of glycolytic rates, albeit intracellular glucose concentrations remain constant (Neves and Francois, 1992). Decreasing of glycolytic rates corresponded to lower activity of the Ras/cAMP pathway upon thermal shock, which favored trehalose synthesis in detriment of glucose catabolism and cells growth (Shin et al., 1987; Neves and Francois, 1992; Piper, 1993; Tokiwa et al., 1994). After recovering homeostasis, cells increase Ras/cAMP pathway activity and glycolytic fluxes (Piper, 1993). Both circumstances seem to regulate cyclins activity (CLN1, CLN2, and CLN3) and transcription of *CLN3*, which are required for cell cycle progression at the START point in G1 phase (Tokiwa et al., 1994; Shi and Tu, 2013), following bud formation. It was recently shown that accumulation of acetyl-CoA, a central metabolite from glucose catabolism, triggers histone acetylation and transcription of *CLN3* (Shi and Tu, 2013).

During the acquisition of thermotolerance, yeast cells also change the lipid composition of cellular membrane. Temperature increment caused the increase of saturations and length of FAs as well as a reduction of FA composition in membranes (Suutari et al., 1990, 1997). The synthesis of long-chain bases (LCBs), which are important for membrane fluidity and dynamics, and with possible role in the regulation of signal transduction pathways, also increased (Dickson et al., 1997; Jenkins et al., 1997). Changes in the synthesis of these lipids and some sterols upon temperature increase suggest that pathways supporting signaling networks of the cell wall integrity are involved in heat-shock response (Kamada et al., 1995; Verna et al., 1997; Imazu and Sakurai, 2005). Overexpression of genes coding for antioxidants and enzymes involved in carbon metabolism mediated by the stress-responsive transcription factors (TFs) MSN2 and MSN4, but the Ras/cAMP/PKA signaling pathway cAMP had a negative effect on the induction of the MSN2/MSN4 regulon (Boy-Marcotte et al., 1999; Imazu and Sakurai, 2005). Hence, the former could mainly occur in the precondition effect of trehalose accumulation.

Collectively, these results suggest a toxicity model in which inhibitory conditions associated with ethanol production processes mainly affect cellular membrane concomitant with exchange reactions between the intracellular and extracellular environment, e.g., protons/ions exchange. Accumulation of toxic chemicals through the pretreatment and fermentation operations eventually exacerbates energy requirements and the cell's effort to maintain gradients and to continue the excretion of toxic chemicals. Although mild pretreatment operations or incorporation of detoxification processes reduce the concentration of toxic compounds, such options have to be carefully considered as it may increase the costs and energy consumption. Cellular gradients can also be maintained by increasing medium pH or supplementing with specific salts. Adaptation of the yeast to process conditions through heritable modifications is, however, the



ideal solution, and for this, it is necessary to understand the basic molecular mechanisms underlying the stress response in yeast.

## THE CELLULAR STRESS RESPONSE IN YEAST

Organisms have developed strategies to mount stress responses to recover the constancy of internal state (homeostasis) upon being exposed to environmental changes (Tagkopoulos et al., 2008; Mitchell et al., 2009). These responses are associated to damage in cellular macromolecules and/or redox potentials which disrupt cellular functions (Kültz, 2005; Gibney et al., 2013). Thus, the CSR is universal and have a define set of targeted cellular functions including cell cycle control, protein chaperoning and repair, DNA and chromatin stabilization and repair, cellular membrane stabilization and repair, removal of damaged proteins, and some aspects of metabolism (Kültz, 2005). This assumption raised from the analysis of around 300 highly conserved proteins among different organisms including human, yeast, eubacteria,

and archaea, from which more than 44 proteins change their abundance upon stress exposition (Kültz, 2003). Here, it has been pointed out, based on recent results, that most of the inhibitory mechanisms target cellular membrane, redox potentials, and exchange reactions functions.

The results from transcriptomic and proteomic analyses altogether suggest that CSRs in *S. cerevisiae* overlap at specific stress conditions, whereas some responses are stress specific (Figure 2). Remarkably, this flexibility allows the coordination of stress responses according to a serial of ordered events, which are naturally organized according to its habitat. For example, yeast mounts a stress response to heat, which also serves to tolerate the stress imposed by ethanol and an oxidative environment. This could be a consequence of the domestication of yeast as these stresses appear in this order during the wine production processes (Mitchell et al., 2009). Yeast also triggers a general stress response to survive exposure to several different types of stress. Thus, when *S. cerevisiae* is exposed to environmental perturbations including temperature increase, nutrient depletion, addition of oxygen peroxide, starvation and stationary phase,

DNA damaging agents, and hyperosmotic stress among others, a set of around 900 genes showed similar changes in their expression (Gasch et al., 2000). Functional analysis of these genes showed similar targeted functions than those found in proteome analysis (Kültz, 2003, 2005). Differentially expressed genes include those encoding for proteins involved in the RAS-cAMP signaling pathway, which regulate cell metabolism and cell cycle progression in response to nutrient availability (Broach, 1991; Gasch and Werner-Washburne, 2002). Simultaneously, transcription of genes involved in the CSR is also modified by the action of a set of TFs including MSN2/4, YAP1, HSF1, RLM1, and SWI6. Activation of these TFs occurs through phosphorylation cascades triggered by structural changes in proteins located in the cellular membrane; this seems to be the major target for stress agents and the origin of signaling pathways for response to stress (Figure 2).

Reprogramming of gene expression in yeast is mainly governed by the general stress response TFs MSN2 and/or MSN4 targeting around 180 genes in response to thermal stress and oxygen peroxide (Gasch et al., 2000). MSN2/MSN4 also induces similar stress responses when yeast cells are exposed to other stresses, suggesting that these TFs induce a general response to various environmental changes (Causton et al., 2001). With regard to the cellular response to oxidative stress, the yeast basic leucine zipper (bZIP) TF YAP1 is mainly in charge of regulating the expression of genes associated to this stress (Temple et al., 2005), as well as in response to xenobiotic insults, including drugs and heavy metals (Lushchak, 2011). Remarkably, transcription of some genes targeted by oxidative stress is also activated by TFs MSN2/MSN4 in response to heat-shock (Gasch et al., 2000). Nonetheless, this response is mainly controlled by the heat-shock TF HSF1, which directs the expression of around 150 genes (Hahn et al., 2004). This TF also triggers gene expression changes upon starvation (Hahn and Thiele, 2004), as well as controls the expression of genes associated to life span extension (Shama et al., 1998). Furthermore, it regulates cellular wall remodeling in response to thermal and oxidative stresses (Imazu and Sakurai, 2005; Yamamoto et al., 2007). Some other relevant aspects of particular regulation of CSRs to stressors found in ethanol production processes are given below.

## Yeast Responses to Inhibitory Compounds

Metabolic and molecular responses of yeast exposition to inhibitory compounds, such as furfural and HMF, caused changes in expression of around 886 genes (Ask et al., 2013a). Functional examination of proteomic and transcriptomic analyses showed that genes involved in redox balance, oxidative and salt stress as well as the TFs, MSN2/MSN4, YAP1, and HSF1 were mostly involved in stress responses and specifically, overexpression of YAP1 and MSN2 were related to the increase of yeast tolerance to furfural and HMF (Lin et al., 2009; Sasano et al., 2012). Activity of the MAPK signaling pathway of the yeast response to cell wall integrity was also found to increase yeast tolerance to HMF (Larsson et al., 1999). In agreement with the transcriptional analysis in reference (Dickson et al., 1997), proteomic analysis also showed that redox and energy metabolism are significantly targeted by the stress response in yeast exposed to hydrolyzates

containing furans, acids, and/or phenolics (Lin et al., 2009; Lv et al., 2014). In the last study, the authors observed differential expression of around 200 genes, a number similar to the 103 and 227 differentially expressed genes observed from yeast exposition to furfural and acetate, respectively (Li and Yuan, 2010). In this study, it was found that tolerance to furfural also required the overexpression of genes involved in the oxidative stress response, such as *SRX1*, *CTA1*, and *GRX5* as well as the *HSP78*, which encodes a mitochondrial chaperone needed for the thermotolerance of this organelle (Heer and Sauer, 2008). In addition, overexpression of some genes related to the lipid and carbohydrate metabolism have been observed within these genes. Interestingly, proteins involved in the TCA cycle were upregulated whereas enzymes of glycerol synthesis were downregulated (Lin et al., 2009). The later results strongly suggest an increment of NADH demand for furans conversion to alcohols and that this reducing power is generated in TCA cycle.

## Yeast Responses to Thermal, Ethanol, and Osmotic Stresses

In response to heat, *S. cerevisiae* typically shows transcriptional changes in genes encoding metabolic enzymes (e.g., hexokinase, glyceraldehyde-3-phosphate dehydrogenase, glucose-6-phosphate dehydrogenase, isocitrate dehydrogenase, and ADH), antioxidant enzymes (e.g., thioredoxin 3, thioredoxin reductase, and porin), molecular chaperones and their cofactors (e.g., *HSP104*, *HSP82*, *HSP60*, *HSP42*, *HSP30*, *HSP26*, *CPR1*, *STI1*, and *ZPR1*), and the TFs (e.g., *HSF1*, *MSN2/4*, and *YAP1*), among others (Lindquist, 1986; Piper, 1993; Kim et al., 2013). Most of these genes also change their expression in response to ethanol and high osmolarity (Gasch et al., 2000; Gasch and Werner-Washburne, 2002). However, stress-specific changes in gene expression also responded solely to either YAP1 or MSN2/4 (Gasch et al., 2000). In addition to the implication of TFs in the regulation of gene transcription upon exposure to different types of stress, another intriguing fact is that the disaggregase protein HSP104 and the negative regulator of the H(+)-ATPase, the HSP30 are overexpressed upon exposition to ethanol, heat, and high osmolarity (Sanchez et al., 1992; Piper et al., 1997; Kültz, 2005). These proteins are implicated in the recovery of aggregated proteins and prevent the cells from excessive energy consumption.

Cellular signaling networks of growth and stress response are antagonist. The RAS-PKA pathway, which regulates yeast proliferation in response to nutritional sensing, negatively regulates the activity of the stress-responsive elements (STRE) and the heat-shock elements (HSE) targeted by both RIM15 and MSN2/4, and HSF1 and MSN2/4, respectively (Roosen et al., 2005). Thus, high activity of the RAS-PKA pathway caused by deletion of the *BCY1* is in detriment of stress responses, whereas deletion of *RAS2* increased yeast resistance to various stresses except high temperature and osmolarity (Ruis and Schüller, 1995) – this is due to the fact that trehalose metabolism is regulated by NTH1, which is probably activated by the RAS-PKA pathway. High activity of this pathway reduces RIM15 activity, which controls the entry into the G0 phase of cell cycle in response to glucose limitation at the diauxic shift. Its regulon includes gene clusters implicated in



the adaptation to respiratory growth, including oxidative stress genes (Cameroni et al., 2004). TFs RIM15, GIS1, and MSN2/4 exerts control on genes required for adaptation to oxidative and thermal stress (Cameroni et al., 2004). High RAS-PKA activity favors the activity of SCH9 kinase, which regulates ribosome biogenesis and translation initiation. This is a major target of TORC1 phosphorylation cascade, transiently reduced upon application of osmotic, oxidative, or thermal stress (Urban et al., 2007). Under favorable conditions, TORC1 promotes growth and antagonizes stress response programs (De Virgilio et al., 1994; Jacinto and Hall, 2003). Thus, TORC1 activity is reduced upon stress apparently by its sequestration in granules (Takahara and Maeda, 2012).

The RAS-PKA pathway also connects with the cell wall damage response. PCK1 and the upstream protein elements ROM2 and MTL1 of the PKC1-MAPK cell integrity pathway are needed for actin organization, and required for cellular responses to oxidative, osmotic, and heat stresses (Kamada et al., 1995; Vilella et al., 2005). More evidences on this fact were provided by two different research groups which discovered that the sensitivity to high osmolality in the HOG-MAPK pathway mutants was reduced at elevated temperature, suggesting that the activation of the cell wall integrity pathway is mainly due to increased temperatures (Alonso-Monge et al., 2001; Wojda et al., 2003). These two pathways and the SVG pathway ensure a proper response of cell wall integrity. The latter is activated by the SHO1 sensor, which also regulates HOG signaling (Figure 2). Furthermore, it was found that membrane sensors WSC1, WSC2, and WSC3 restored the thermo-sensible phenotype of *RAS1* mutants – WSC triple mutants did not growth at 37°C (Verna et al., 1997), which is another evidence for the connection between RAS-PKA signaling cascade and the cell integrity pathway. Signal transduction of this pathway begins with the cellular membrane proteins WSC1-3, MID2, and MTL1, among others (Rodicio and Heinisch, 2010). Finalizing with the phosphorylation of the TF SWI6 leading its localization into the nucleus required for the unfolded protein response (Scrimale et al., 2009).

Thermal stress also has an important effect on the metabolic responses – e.g., glucose and oxygen consumption rates and biomass yields. In nitrogen-limited chemostats, glucose consumption rate increased up to 1.8 times at 38°C compared to 30°C (Postmus et al., 2008). Besides ethanol production rate increased 1.7 times, its yield decreased 0.6 times. Furthermore, in the cultivations at 38°C, glycerol was accumulated at 1.3 mmol g<sub>DCW</sub><sup>-1</sup> h<sup>-1</sup> but no accumulation was observed at 30°C (Postmus et al., 2012). Despite oxygen uptake rate increase 1.1 times at high temperature, respiratory quotients (RQs) of 2.6 and 3.8 were calculated for the fermentations at low and high temperatures, respectively (Postmus et al., 2012). In the same work, a drastic drop of biomass yield was observed in cultivations grown at high temperature (38°C) as compared to the cultures developed at low temperature (Postmus et al., 2008). This behavior correlated with an increased flux of glycerol and ethanol at 38°C – these were not observed at 30°C. In both cases, oxygen consumption rate slightly increased suggesting that reducing power produced in glycolysis is balanced by glycerol production and interrupted electron transport chain.

These results altogether show the complexity of cell stress responses and the difficulties for generating complex thermotolerant phenotypes. Therefore, selection of thermotolerant microorganisms from harsh environments similar to those found in ethanol production process is still a recurrent option. However, these microorganisms will eventually be useless when process conditions change – this will be especially true in the foundation of lignocellulosic biorefineries. Thus, a rational cell design based on knowledge of cell responses will enable the design of generic cell factories that can be used in several different processes. One must be aware that there would be physical components that limit biological augmentation. In this case, synthetic biology approaches (Alper and Stephanopoulos, 2009) and utilization of additional operations to remove toxic molecules can be of interest. In the last section, we review the current strategies for rational improving of yeast tolerance to ethanol production processes.

## IMPROVING YEAST TOLERANCE TO INHIBITORY CONDITIONS FOUND IN ETHANOL PRODUCTION PROCESS

Some of the methods for increasing yeast tolerance to harmful conditions found in ethanol production processes are summarized in Table 2. These methods include the adaptive laboratory evolution (ALE), which is performed by serial dilution of microbial population in fresh media, maintaining or increasing the intensity of the stress (Elena and Lenski, 2003). The main advantage of this method is that increased fitness can be followed during the evolution, and populations can be screened for a strain with a useful phenotype – e.g., improved growth or increased glucose consumption. When combined with partial or complete genome sequence of isolated strains/populations, as well as genome level analysis of gene expression and metabolic fluxes, this procedure is very powerful to get basic knowledge about cell strategies that arise with the better performance (Hong et al., 2011; Caspeta et al., 2014b).

Evolution of linear DNA fragments upon recombination of blocks of sequences rather than point mutagenesis alone has shown to be more important during evolution. The DNA shuffling technology is a procedure for rapid propagation of beneficial mutations in a direct evolution experiment (Stemmer, 1994). This is based on repeated cycles of point mutagenesis, recombination, and selection allowing molecular evolution of complex sequences, through increasing the size of DNA library (Zhang et al., 2002). In combination with cellular mating, this technology has led the generation of cells resistance to ethanol production processes (Pinel et al., 2011).

Despite changes on expression of a single gene have given good results in generating the tolerant phenotype (Caspeta et al., 2014b; Lam et al., 2014), this does not typically occurs since the tolerance to stressors requires changes of expression for thousand genes (see The Cellular Stress Response in Yeast). Therefore, the engineering of global transcription machinery has been developed to generate TFs that may lead with a proper reprogramming of gene transcription network, which arises with the desired tolerant

**TABLE 2 | Some examples of the strategies to improve yeast stress tolerance to inhibitory conditions during the conversion of lignocellulosic biomass to ethanol.**

Molecular strategies	Specific cases	Reference
Adaptive evolution	The industrial <i>S. cerevisiae</i> ethanol red was subjected to a long-term adaptive laboratory evolution. The resultant strain was able to grow and produce ethanol using non-detoxified spruce hydrolyzates Using adaptive laboratory evolution/visualizing evolution at real time, seven strains were isolated due to their improved tolerance to lignocellulosic biomass hydrolyzates <i>S. cerevisiae</i> strain CENPK113-7D was evolved in laboratory and seven thermotolerant strains were isolated. These strains were able to grow at 40°C under fully aerobic conditions with improved kinetic parameters as compared to the parental strain	Wallace-Salinas and Gorwa-Grauslund (2013) Almario et al. (2013) Caspeta et al. (2014b)
Reprogramming gene expression	Mutagenesis of the transcription factor SPT15 allows to increase osmotic and ethanol tolerance, improving ethanol production	Alper et al. (2006)
Direct evolution – DNA shuffling technology	The strain <i>S. cerevisiae</i> R57 was isolated after five rounds of DNA evolution by DNA shuffling. It was able to survive, grow, and produce ethanol using a substrate, hardwood spent sulfite liquor. This strain also increased its viability in presence of salt, peroxide, acetic acid, and sorbitol	Pinel et al. (2011, 2015)
Random mutations (by chemical agents or UV)	Three <i>S. cerevisiae</i> strains with improved tolerance to vanillin, furfural, and acetic acid were generated by random mutagenesis with ethyl methane sulfonate coupled to an adaptive laboratory evolution strategy	Shen et al. (2014)
Heterologous expression	The laccase I from <i>Trametes sp.</i> was cloned in yeast. The laccase-1 displaying yeast has oxidation activity and was effective during pretreatment for ethanol fermentation	Nakanishi et al. (2012)

phenotype (Alper and Stephanopoulos, 2009). This method consists on the mutagenesis of TFs acting with a desired promoter sequences – e.g., the TATA-binding, and the selection of dominant mutations conferring the desired tolerant phenotype.

Random mutagenesis with chemical or physical agents, for example, the dimethyl sulfate (DMS) and UV radiation, has been used for long time to generate populations with a set of mutations from which the useful ones are selected from experiments with the desired environmental pressure. This method has been used in combination with cellular shorting procedures to analyze thousands of phenotypes and came up with the most desired one (Huang et al., 2015). Despite this method being useful for generating tolerant phenotypes, it rarely permits the analysis of mutations that arise with the desired phenotype.

## Increasing Ethanol Tolerance

Despite *S. cerevisiae* showing high ethanol tolerance, there have been many efforts to enhance this trait and generate strains tolerant to higher concentrations; here are some of the recent advances. Comparison of gene expression among tolerant and non-tolerant strains has served to recognize target genes involved in ethanol tolerance. Some genes involved in this feature are the global TF MSN2, some genes of the cAMP-PKA signaling pathway, genes related to the cellular wall integrity, and some genes encoding enzymes of lipids and carbohydrates metabolism (Lewis et al., 2010). It was shown recently that the manipulation of ions transport systems can also improve ethanol tolerance. For instance, changing potassium ion and proton electrochemical forces can improve yeast tolerance to ethanol (Lam et al., 2014). Overexpression of the *TRK1* gene, a member of the potassium transport system, and the H(+)-ATPase gene, *PMA1*, in laboratory strains increased ethanol production by around 30% respect to the laboratory strain S288C and by 10% compared to industrial strains (Lam et al., 2014). In contrast to those findings, thermally evolved *S. cerevisiae* strains, which showed slight increase of

ethanol tolerance, did not overexpress *PMA1* (Caspeta et al., 2014b). The negative regulator of the H(+)-ATPase pump, the gene *HSP30*, however, increased upon thermal stress (Piper et al., 1997; Meena et al., 2011), suggesting that thermal adaptation may optimize ATP usage for proton excretion, thus decreasing energy for maintenance. Thereafter, electrical potential and proton fluxes can decrease free energy of ATP hydrolysis for proton export (Maris et al., 2004), enhancing the resistance to alcohols (Lam et al., 2014).

Transcription reprogramming of yeast gene expression using the global transcription machinery engineering approach led with higher ethanol resistance. The mutagenesis of the TF SPT15 allowed the selection of the SPT15-300 TF with a mutation in the phenylalanine (Phe<sup>177</sup> Ser) as the dominant mutation which provided increased tolerance to elevated concentrations of glucose and ethanol, as well as improved ethanol production (Alper et al., 2006).

## Increasing Tolerance to Toxic Compounds

Adaptive laboratory evolution has been successfully used for selection of yeast strains tolerant to lignocellulose hydrolyzates containing furfural, HMF, and acetate (Liu et al., 2005; Keating et al., 2006; Heer and Sauer, 2008). Evolution of yeast populations in synthetic medium containing 3 mM furfural resulted in the selection of tolerant strains after 300 generations (Heer and Sauer, 2008). These strains reduced the lag-phase of growth suggesting that furfural conversion to its alcohol is the main factor for improving the tolerance. In agreement with this, the evolution of the industrial yeast strain TMB3400 in synthetic mixtures of sugars supplemented with furfural, HMF, and acetic acid showed faster consumption of these inhibitors (Keating et al., 2006). Furthermore, the conversion of furfural to furfuryl alcohol at significantly higher rates was the solution of evolved *S. cerevisiae* and *Pichia pastoris* strains to tolerate these chemicals (Liu et al., 2005). These results suggest that detoxification of furfural and

HMF can be carried out in place with yeast strains having higher ability to convert such toxic molecules.

Evolution of the industrial strain ethanol red of *S. cerevisiae* in non-detoxified spruce hydrolyzate in combination to high temperature (39°C) resulted in the selection of strains capable to convert spruce hydrolyzates into ethanol with high efficiency (Wallace-Salinas and Gorwa-Grauslund, 2013). Contrary to the resistance in evolved strains selected with furfural and HMF alone, the superior phenotype of the evolved ethanol red strains did not rely on higher reductase activities for furfural conversion, but rather on a higher thermotolerance. Different results were also observed in tolerant yeast strains obtained from evolutionary engineering using genome-shuffling technology based on large-scale population with cross-mating to generate tolerance to spent sulfite liquor (SSL) (Pinel et al., 2011). These strains were also more tolerant to higher osmolarities, elevated ethanol concentrations, and higher amounts of acetic acid than the parental strain.

Studies based on the change in gene expression using microarrays have led to the identification of redox balance and energy state of the cells as the major drivers to generate tolerance to furfural and HMF (Petersson et al., 2006; Ask et al., 2013a). From the 15 reductases which overexpression were found to improve tolerance, the overexpression of three candidate genes raised with the recognition of *ADH6* as one of the major contributors for tolerance to HMF in aerobic and anaerobic conditions (Petersson et al., 2006). It has been also demonstrated that tolerance to furfural can be increased by the overexpression of *ADH7*, the ORF *YKL071W*, and *ARI1* genes, which encode are reductases involved in furfural reduction (Heer et al., 2009; Sehnem et al., 2013). Combining the overexpression of the ADH *ADH1* with the transaldolase *TAL1* in recombinant xylose-fermenting *S. cerevisiae* improves ethanol production from lignocellulosic hydrolyzates. Most of the tolerant strains generated by these means increase furfural and HMF conversion to their less toxic alcohols. This strategy has been also effective in *E. coli*, in which the overexpression of reductases *YGHD* and *DKGA*, having NADPH-dependent furfural reductase activity, increases furfural tolerance (Miller et al., 2009).

Besides the overexpression of *TAL1*, the overexpression of some genes of the pentose phosphate pathway also increases yeast tolerance to furfural. Among them, the overexpression of *ZWF1*, *GND1*, or *RPE1* induced tolerance to furfural at concentrations that are normally toxic in the wild-type strain (Gorsich et al., 2006). These results were similar to those observed when the xylose reductase and xylitol dehydrogenase from *P. pastoris* were overexpressed in combination with overexpression of the endogenous xylulose kinase of *S. cerevisiae* (Almeida et al., 2008). On the other hand, the overexpression of *YAP1* activated the transcription of catalases genes *CTA1* and *CTT1*, enhancing the tolerance to furfural but not to HMF (Kim and Hahn, 2013), suggesting that rapid furfural consumption is associated to accumulation to reactive oxygen species.

The evolutionary engineering through genome-shuffling technology was used to increase yeast tolerance to hardwood SSL (Pinel et al., 2011). Using RNA-seq gene expression analysis, these authors found that the products of the genes *UBP7* and *ART5* (both related to ubiquitin-mediated proteolysis), *NRG1* (a

stress-response transcriptional repressor), and *GDH1* (a NADPH-dependent glutamate dehydrogenase), play an important role in the tolerance to these hydrolyzates (Pinel et al., 2015). The genome-shuffling technology method was also used to increase tolerance to a combination of heat, acetic acid, and furfural stresses (Lu et al., 2012). The resulted strains showed tolerance to 0.55% (v/v) acetic acid and 0.3% (v/v) furfural at 40°C.

Tolerance to phenolics can be tackled by the expression of extracellular heterologous laccases (Lee et al., 2012). This trait can be also enhanced by heterologous expression of the gene encoding the phenyl acrylic-acid decarboxylase (PSP1), which catalyzes the decarboxylation of aromatic carboxylic acids into the corresponding vinyl derivatives (Richard et al., 2015). Overexpression of multidrug efflux pump genes *ATR1* and *FLR1*, and the TF *YAP1* also resulted in yeast resistance to coniferyl aldehyde and HMF (Alriksson et al., 2010). Tolerance to vanillin and 39°C were induced after several rounds of mutagenesis in hydrolyzates containing vanillin (Kumari and Pramanik, 2012). Chemical mutagenesis coupled with ALE using continuous cultivation in 60% (v/v) non-detoxified hydrolyzate liquor from steam-pretreated lignocellulose was successfully used to select yeast strains with improved capacity to ferment xylose from lignocellulose hydrolyzates (Smith et al., 2014). Since many of the toxic compounds affect the membrane potential, the addition of spermidine, which synchronize  $\text{Ca}^{2+}$ ,  $\text{Na}^{+}$ ,  $\text{K}^{+}$ , and ATPase has also proven to induce tolerance, after disruption of the spermidine metabolism genes *OAZ1* coding for an ornithine decarboxylase (ODC) enzyme and *TPO1* coding for the polyamine transport protein (Kim et al., 2015). Changes in ergosterol composition have shown to improve tolerance to vanillin in strains overexpressing the ergosterol synthesis genes *ERG28*, *HMG1*, *MCR1*, *ERG5*, and *ERG7* (Endo et al., 2009).

## Increasing Tolerance to High Temperature and Elevated Osmolality

Adaptive laboratory evolution has been successful for selecting thermotolerant *S. cerevisiae* strains (Yona et al., 2012; Caspeta et al., 2014b). After an evolution period of 450 generations, thermotolerant yeast populations were isolated from experiments performed at 39°C (Yona et al., 2012). These populations showed a duplication in chromosome number III (ChIII) and overexpression of related genes. However, only the overexpression of some genes found in this chromosome including the TF *HCM1* and the protease *RRT12* reproduced a significant fraction of the thermotolerant phenotype in the parental strain. A similar result was found in thermotolerant yeast strains isolated from ALE experiments to 39.5°C (Caspeta et al., 2014b). In this work, a partial duplication of ChIII containing the *HCM1* gene was found. Since duplication of ChIII was lost in evolved strains, this suggests that chromosomal duplications are a temporal solution to stress (Yona et al., 2012).

Adaptive laboratory evolution experiments have also used to generate tolerance to high pH, which induced the duplication of chromosome number five (Yona et al., 2012). Remarkably, chromosomal duplications only appear in diploid cells since haploid *S. cerevisiae* populations showed segmental duplications only. In



these strains, a nonsense mutation of *ERG3* proportionated 80% of the thermotolerant phenotype (Caspeta et al., 2014b). This result and the fact that ethanol tolerance can be achieved by just one overexpression suggest that complex tolerant phenotypes can be achievable by just one mutation. Remarkably, this mutation changed cellular membrane properties.

Genome-shuffling technology was used to improve yeast performance in high gravity fermentations (Liu et al., 2011). The resulted strains derived from a diploid STE2/STE2 (receptor for alpha-factor pheromone) strain increased tolerance to high osmolality and elevated ethanol concentrations. This method was also used to generate sexual and asexual populations of *S. cerevisiae* resistant to very high gravity fermentations, elevated temperature, and high glucose concentrations (Hou, 2010). In mutants of the gene *GPD2* encoding glycerol 3-phosphate dehydrogenase subjected to three rounds of genome shuffling, a population of strains producing lower amounts of glycerol and improved tolerance to ethanol and high osmolality were able to be selected (Tao et al., 2012). These strains showed changes in FAs composition and higher accumulation of trehalose. A remarkable application of the genome-shuffling technology was the generation of both thermotolerance and ethanol tolerance in the industrial yeast strain SM-3, which were used to ferment syrups with 20% (w/v) glucose at 45°C and resists 9.5% (w/v) ethanol (Shi et al., 2009).

## CONCLUDING REMARKS

One of the major challenges for economic conversion of lignocellulose to fuel ethanol is to generate robust *S. cerevisiae* strains able to cope with inhibitory conditions while keeping proper catalytic functions for raw material conversion to ethanol. Major inhibitory conditions found in the unit operations required for the conversion processes include the accumulation of toxic chemicals generated during lignocellulose pretreatment and sugar fermentation, the high temperature that accompanied simultaneous saccharification and fermentation, and the very high osmolality and elevated solids loadings at the beginning of the fermentation. Since unification of these unit operations is desirable to reduce production costs and energy utilization, it can be expected that yeast cells will be simultaneously exposed to most of these inhibitory conditions.

## REFERENCES

- Adeboye, P. T., Bettiga, M., and Olsson, L. (2014). The chemical nature of phenolic compounds determines their toxicity and induces distinct physiological responses in *Saccharomyces cerevisiae* in lignocellulose hydrolysates. *AMB Express* 4, 46–46. doi:10.1186/s13568-014-0046-7
- Aguilera, F., Peinado, R. A., Millán, C., Ortega, J. M., and Mauricio, J. C. (2006). Relationship between ethanol tolerance, H<sup>+</sup>-ATPase activity and the lipid composition of the plasma membrane in different wine yeast strains. *Int. J. Food Microbiol.* 110, 34–42. doi:10.1016/j.jfoodmicro.2006.02.002
- Allen, S., Clark, W., Mccaffery, J. M., Cai, Z., Lancot, A., Slininger, P., et al. (2010). Furfural induces reactive oxygen species accumulation and cellular damage in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* 3, 2. doi:10.1186/1754-6834-3-2
- Almario, M. P., Reyes, L. H., and Kao, K. C. (2013). Evolutionary engineering of *Saccharomyces cerevisiae* for enhanced tolerance to hydrolysates of lignocellulosic biomass. *Biotechnol. Bioengin.* 110, 2616–2623. doi:10.1002/bit.24938

Since cellular macromolecules and metabolism have evolved to sustain optimal growth rates at the prevailed natural conditions, mainly preserving genetic information and proteins/membrane functional structures, the generation of complex tolerant phenotypes for the ethanol industry will be further generated on the bases of the functions targeted by the stressors. The summarized results altogether show that major targets include cellular membrane, redox and ionic potentials, and energy metabolism, as well as protein structure – the latter of apparently minor relevance.

To establish metabolic engineering strategies for increasing yeast tolerance, it is suggested to consider the route and regulation of molecular responses following sensing, signal transduction, signal integration, and execution of cellular functions in response to environmental stresses. Results from systems biology and -omics analyses, as well as from traditional data mining, point out the relevance of the cross-regulation between the routes of yeast responses according to the different types of stress. This is part of the elasticity of cellular stress-signaling network, which is advantageous during evolutionary adaptation and in the generation of resistance to the multiple stresses found in ethanol production process.

In summary, the multiple technologies for the generation of numerous mutations, high-throughput screening, acceleration of cells adaptation and selection, laboratory evolution and engineering of TFs, and the new tools for controlling gene expression are accelerating the accumulation of basic information of CSRs, and the generation of yeast cells with desirable processing characteristics including better performance in the inhibitory conditions found in lignocellulosic ethanol production processes.

## ACKNOWLEDGMENTS

LC thanks the Secretaría de Educación Pública (SEP) which financed his integration into the Autonomous University of Morelos State and covered publication fees through the Programa para el Desarrollo Profesional Docente (PRODEP) (grant no. DSA/103.5/14/10703). TC thanks the National Council of Science (CONACyT) for the postdoctoral scholarship. JN thanks Novo Nordisk Foundation, the Knut och Alice Wallenbergs Stiftelse, and Vetenskapsrådet for the funding received to perform work related to the topics covered in this review.

- Almeida, J., Modig, T., Roder, A., Liden, G., and Gorwa-Grauslund, M.-F. (2008). *Pichia stipitis* xylose reductase helps detoxifying lignocellulosic hydrolysate by reducing 5-hydroxymethyl-furfural (HMF). *Biotechnol. Biofuels* 1, 12. doi:10.1186/1754-6834-1-12
- Alonso-Monge, R., Real, E., Wojda, I., Bebelman, J.-P., Mager, W. H., and Siderius, M. (2001). Hyperosmotic stress response and regulation of cell wall integrity in *Saccharomyces cerevisiae* share common functional aspects. *Mol. Microbiol.* 41, 717–730. doi:10.1046/j.1365-2958.2001.02549.x
- Alper, H., Moxley, J., Nevoigt, E., Fink, G. R., and Stephanopoulos, G. (2006). Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* 314, 1565–1568. doi:10.1126/science.1131969
- Alper, H., and Stephanopoulos, G. (2009). Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? *Nat. Rev. Microbiol.* 7, 715–723. doi:10.1038/nrmicro2186
- Alriksson, B., Horváth, I. S., and Jönsson, L. J. (2010). Overexpression of *Saccharomyces cerevisiae* transcription factor and multidrug resistance genes



- conveys enhanced resistance to lignocellulose-derived fermentation inhibitors. *Process Biochem.* 45, 264–271. doi:10.1016/j.procbio.2009.09.016
- Ando, S., Arai, I., Kiyoto, K., and Hanai, S. (1986). Identification of aromatic monomers in steam-exploded poplar and their influences on ethanol fermentation by *Saccharomyces cerevisiae*. *J. Ferment. Technol.* 64, 567–570. doi:10.1016/0385-6380(86)90084-1
- Ando, T. (1966). A nuclease specific for heat-denatured DNA is isolated from a product of *Aspergillus oryzae*. *Biochim. Biophys. Acta* 114, 158–168. doi:10.1016/0005-2787(66)90263-2
- Ask, M., Bettiga, M., Duraiswamy, V., and Olsson, L. (2013a). Pulsed addition of HMF and furfural to batch-grown xylose-utilizing *Saccharomyces cerevisiae* results in different physiological responses in glucose and xylose consumption phase. *Biotechnol. Biofuels* 6, 181. doi:10.1186/1754-6834-6-181
- Ask, M., Bettiga, M., Mapelli, V., and Olsson, L. (2013b). The influence of HMF and furfural on redox-balance and energy-state of xylose-utilizing *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* 6, 22. doi:10.1186/1754-6834-6-22
- Augustin, H. W., Kopperschlager, G., Steffen, H., and Hofmann, E. (1965). Hexokinase as limiting factor of anaerobic glucose consumption of *Saccharomyces carlsbergensis* NCYC74. *Biochim. Biophys. Acta* 110, 437–439. doi:10.1016/S0926-6593(65)80055-8
- Ballesteros, I., Ballesteros, M., Cabañas, A., Carrasco, J., Martín, C., Negro, M. J., et al. (1991). Selection of thermotolerant yeasts for simultaneous saccharification and fermentation (SSF) of cellulose to ethanol. *Appl. Biochem. Biotechnol.* 2, 307–315. doi:10.1007/bf02922610
- Banerjee, N., Bhatnagar, R., and Viswanathan, L. (1981). Inhibition of glycolysis by furfural in *Saccharomyces cerevisiae*. *Eur. J. Appl. Microbiol. Biotechnol.* 11, 226–228. doi:10.1007/bf00505872
- Boy-Marcotte, E., Lagniel, G., Perrot, M., Bussereau, F., Boudsocq, A., Jacquet, M., et al. (1999). The heat shock response in yeast: differential regulations and contributions of the Msn2p/Msn4p and Hsf1p regulons. *Mol. Microbiol.* 33, 274–283. doi:10.1046/j.1365-2958.1999.01467.x
- Broach, J. R. (1991). RAS genes in *Saccharomyces cerevisiae*: signal transduction in search of a pathway. *Trends Genet.* 7, 28–33. doi:10.1016/0168-9525(91)90018-L
- Brown, S., and Oliver, S. (1982). The effect of temperature on the ethanol tolerance of the yeast, *Saccharomyces uvarum*. *Biotechnol. Lett.* 4, 269–273. doi:10.1007/bf00132399
- Cameroni, E., Hulo, N., Roosen, J., Winderickx, J., and Virgilio, C. D. (2004). The novel yeast PAS kinase Rim15 orchestrates G0-associated antioxidant defense mechanisms. *Cell Cycle* 3, 460–466. doi:10.4161/cc.3.4.791
- Casal, M., Cardoso, H., and Leao, C. (1996). Mechanisms regulating the transport of acetic acid in *Saccharomyces cerevisiae*. *Microbiology* 142, 1385–1390. doi:10.1099/13500872-142-6-1385
- Caspeta, L., Buijs, N. A. A., and Nielsen, J. (2013). The role of biofuels in the future energy supply. *Energy Environ. Sci.* 6, 1077–1082. doi:10.1039/c3ee24403b
- Caspeta, L., Caro-Bermúdez, M. A., Ponce-Noyola, T., and Martínez, A. (2014a). Enzymatic hydrolysis at high-solids loadings for the conversion of agave bagasse to fuel ethanol. *Appl. Energy* 113, 277–286. doi:10.1016/j.apenergy.2013.07.036
- Caspeta, L., Chen, Y., Ghiaci, P., Feizi, A., Buskov, S., Hallström, B. M., et al. (2014b). Altered sterol composition renders yeast thermotolerant. *Science* 346, 75–78. doi:10.1126/science.1258137
- Caspeta, L., and Nielsen, J. (2013). Economic and environmental impacts of microbial biodiesel. *Nat. Biotechnol.* 31, 789–793. doi:10.1038/nbt.2683
- Causton, H. C., Ren, B., Koh, S. S., Harbison, C. T., Kanin, E., Jennings, E. G., et al. (2001). Remodeling of yeast genome expression in response to environmental changes. *Mol. Biol. Cell* 12, 323–337. doi:10.1091/mbc.12.2.323
- Chen, H., and Jin, S. (2006). Effect of ethanol and yeast on cellulase activity and hydrolysis of crystalline cellulose. *Enzyme Microb. Technol.* 39, 1430–1432. doi:10.1016/j.enzmtec.2006.03.027
- Chi, Z., and Arneborg, N. (1999). Relationship between lipid composition, frequency of ethanol-induced respiratory deficient mutants, and ethanol tolerance in *Saccharomyces cerevisiae*. *J. Appl. Microbiol.* 86, 1047–1052. doi:10.1046/j.1365-2672.1999.00793.x
- Chowdhury, S., Smith, K. W., and Gustin, M. C. (1992). Osmotic stress and the yeast cytoskeleton: phenotype-specific suppression of an actin mutation. *J. Cell Biol.* 118, 561–571. doi:10.1083/jcb.118.3.561
- D'Amore, T., Panchal, C., Russeil, I., and Stewart, G. (1988). Osmotic pressure effects and intracellular accumulation of ethanol in yeast during fermentation. *J. Ind. Microbiol.* 2, 365–372. doi:10.1007/bf01569575
- De Virgilio, C., Hottiger, T., Dominguez, J., Boller, T., and Wiemken, A. (1994). The role of trehalose synthesis for the acquisition of thermotolerance in yeast. *Eur. J. Biochem.* 219, 179–186. doi:10.1111/j.1432-1033.1994.tb19928.x
- Delgenes, J. P., Moletta, R., and Navarro, J. M. (1996). Effects of lignocellulose degradation products on ethanol fermentations of glucose and xylose by *Saccharomyces cerevisiae*, *Zymomonas mobilis*, *Pichia stipitis*, and *Candida shehatae*. *Enzyme Microb. Technol.* 19, 220–225. doi:10.1016/0141-0229(95)00237-5
- Diaz De Villegas, M. E., Villa, P., Guerra, M., Rodriguez, E., Redondo, D., and Martinez, A. (1992). Conversion of furfural into furfuryl alcohol by *Saccharomyces cerevisiae* 354. *Acta Biotechnol.* 12, 351–354. doi:10.1002/abio.370120420
- Dickson, R. C., Nagiec, E. E., Skrzypek, M., Tillman, P., Wells, G. B., and Lester, R. L. (1997). Sphingolipids are potential heat stress signals in *Saccharomyces*. *J. Biol. Chem.* 272, 30196–30200. doi:10.1074/jbc.272.48.30196
- Edgardo, A., Carolina, P., Manuel, R., Juanita, F., and Baeza, J. (2008). Selection of thermotolerant yeast strains *Saccharomyces cerevisiae* for bioethanol production. *Enzyme Microb. Technol.* 43, 120–123. doi:10.1016/j.enzmtec.2008.02.007
- Elena, S. F., and Lenski, R. E. (2003). Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* 4, 457–469. doi:10.1038/nrg1088
- Endo, A., Nakamura, T., and Shima, J. (2009). Involvement of ergosterol in tolerance to vanillin, a potential inhibitor of bioethanol fermentation, in *Saccharomyces cerevisiae*. *FEMS Microbiol. Lett.* 299, 95–99. doi:10.1111/j.1574-6968.2009.01733.x
- Field, S., Ryden, P., Wilson, D., James, S., Roberts, I., Richardson, D., et al. (2015). Identification of furfural resistant strains of *Saccharomyces cerevisiae* and *Saccharomyces paradoxus* from a collection of environmental and industrial isolates. *Biotechnol. Biofuels* 8, 33. doi:10.1186/s13068-015-0217-z
- Gao, C., and Fleet, G. H. (1988). The effects of temperature and pH on the ethanol tolerance of the wine yeasts, *Saccharomyces cerevisiae*, *Candida stellata* and *Kloeckera apiculata*. *J. Appl. Bacteriol.* 65, 405–409. doi:10.1111/j.1365-2672.1988.tb01909.x
- Garay-Arroyo, A., Covarrubias, A. A., Clark, I., Niño, I., Gosset, G., and Martínez, A. (2004). Response to different environmental stress conditions of industrial and laboratory *Saccharomyces cerevisiae* strains. *Appl. Microbiol. Biotechnol.* 63, 734–741. doi:10.1007/s00253-003-1414-4
- Gasch, A., and Werner-Washburne, M. (2002). The genomics of yeast responses to environmental stress and starvation. *Funct. Integr. Genomics* 2, 181–192. doi:10.1007/s10142-002-0058-2
- Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., et al. (2000). Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* 11, 4241–4257. doi:10.1091/mbc.11.12.4241
- Gibney, P. A., Lu, C., Caudy, A. A., Hess, D. C., and Botstein, D. (2013). Yeast metabolic and signaling genes are required for heat-shock survival and have little overlap with the heat-induced genes. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4393–4402. doi:10.1073/pnas.1318100110
- Gorsich, S. W., Dien, B. S., Nichols, N. N., Slininger, P. J., Liu, Z. L., and Skory, C. D. (2006). Tolerance to furfural-induced stress is associated with pentose phosphate pathway genes ZWF1, GND1, RPE1, and TKL1 in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 71, 339–349. doi:10.1007/s00253-005-0142-3
- Haelssig, J. B., Tremblay, A. Y., and Thibault, J. (2008). Technical and economic considerations for various recovery schemes in ethanol production by fermentation. *Ind. Eng. Chem. Res.* 47, 6185–6191. doi:10.1021/ie0715005
- Hahn, J.-S., Hu, Z., Thiele, D. J., and Iyer, V. R. (2004). Genome-wide analysis of the biology of stress responses through heat shock transcription factor. *Mol. Cell Biol.* 24, 5249–5256. doi:10.1128/mcb.24.12.5249-5256.2004
- Hahn, J.-S., and Thiele, D. J. (2004). Activation of the *Saccharomyces cerevisiae* heat shock transcription factor under glucose starvation conditions by snf1 protein kinase. *J. Biol. Chem.* 279, 5169–5176. doi:10.1074/jbc.M311005200
- Hallsworth, J. E. (1998). Ethanol-induced water stress in yeast. *J. Ferment. Bioeng.* 85, 125–137. doi:10.1016/S0922-338X(97)86756-6
- Hazell, B. W., Nevalainen, H., and Attfield, P. V. (1995). Evidence that the *Saccharomyces cerevisiae* C1F1 (GGS1/TPS1) gene modulates heat shock response positively. *FEBS Lett.* 377, 457–460. doi:10.1016/0014-5793(95)01392-X
- Heer, D., Heine, D., and Sauer, U. (2009). Resistance of *Saccharomyces cerevisiae* to high concentrations of furfural is based on NADPH-dependent reduction by at least two oxidoreductases. *Appl. Environ. Microbiol.* 75, 7631–7638. doi:10.1128/aem.01649-09

- Heer, D., and Sauer, U. (2008). Identification of furfural as a key toxin in lignocellulosic hydrolysates and evolution of a tolerant yeast strain. *Microb. Biotechnol.* 1, 497–506. doi:10.1111/j.1751-7915.2008.00050.x
- Heipieper, H. J., Weber, F. J., Sikkema, J., Keweloh, H., and De Bont, J. A. M. (1994). Mechanisms of resistance of whole cells to toxic organic solvents. *Trends Biotechnol.* 12, 409–415. doi:10.1016/0167-7799(94)90029-9
- Hohmann, S. (2002). Osmotic stress signaling and osmoadaptation in yeasts. *Microbiol. Mol. Biol. Rev.* 66, 300–372. doi:10.1128/mmbr.66.2.300-372.2002
- Holyoak, C. D., Stratford, M., McMullin, Z., Cole, M. B., Crimmins, K., Brown, A. J., et al. (1996). Activity of the plasma membrane H(+)-ATPase and optimal glycolytic flux are required for rapid adaptation and growth of *Saccharomyces cerevisiae* in the presence of the weak-acid preservative sorbic acid. *Appl. Environ. Microbiol.* 62, 3158–3164.
- Hong, K.-K., Vongsangnak, W., Vemuri, G. N., and Nielsen, J. (2011). Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12179–12184. doi:10.1073/pnas.1103219108
- Hottiger, T., Boller, T., and Wiemken, A. (1989). Correlation of trehalose content and heat resistance in yeast mutants altered in the RAS/adenylyl cyclase pathway: is trehalose a thermoprotectant? *FEBS Lett.* 255, 431–434. doi:10.1016/0014-5793(89)81139-1
- Hottiger, T., Schmutz, P., and Wiemken, A. (1987). Heat-induced accumulation and futile cycling of trehalose in *Saccharomyces cerevisiae*. *J. Bacteriol.* 169, 5518–5522.
- Hou, L. (2010). Improved production of ethanol by novel genome shuffling in *Saccharomyces cerevisiae*. *Appl. Biochem. Biotechnol.* 160, 1084–1093. doi:10.1007/s12010-009-8552-9
- Huang, M., Bai, Y., Sjostrom, S. L., Hallström, B. M., Liu, Z., Petranovic, D., et al. (2015). Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast. *Proc. Natl. Acad. Sci. U.S.A.* 112, E4689–E4696. doi:10.1073/pnas.1506460112
- Ibeas, J. I., and Jimenez, J. (1997). Mitochondrial DNA loss caused by ethanol in *Saccharomyces flor* yeasts. *Appl. Environ. Microbiol.* 63, 7–12.
- Imazu, H., and Sakurai, H. (2005). *Saccharomyces cerevisiae* heat shock transcription factor regulates cell wall remodeling in response to heat shock. *Eukaryot. Cell* 4, 1050–1056. doi:10.1128/EC.4.6.1050-1056.2005
- Ingesson, H., Zacchi, G., Yang, B., Esteghlalian, A. R., and Saddler, J. N. (2001). The effect of shaking regime on the rate and extent of enzymatic hydrolysis of cellulose. *J. Biotechnol.* 88, 177–182. doi:10.1016/S0168-1656(01)00273-5
- Jacinto, E., and Hall, M. N. (2003). TOR signalling in bugs, brain and brawn. *Nat. Rev. Mol. Cell Biol.* 4, 117–126. doi:10.1038/nrm1018
- Jenkins, G. M., Richards, A., Wahl, T., Mao, C., Obeid, L., and Hannun, Y. (1997). Involvement of yeast sphingolipids in the heat stress response of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 272, 32566–32572. doi:10.1074/jbc.272.51.32566
- Jones, R. P., and Greenfield, P. F. (1987). Ethanol and the fluidity of the yeast plasma membrane. *Yeast* 3, 223–232. doi:10.1002/yea.320030403
- Jonsson, L., Alriksson, B., and Nilvebrant, N.-O. (2013). Bioconversion of lignocellulose: inhibitors and detoxification. *Biotechnol. Biofuels* 6, 16. doi:10.1186/1754-6834-6-16
- Kajiwar, S., Shirai, A., Fujii, T., Toguri, T., Nakamura, K., and Ohtaguchi, K. (1996). Polyunsaturated fatty acid biosynthesis in *Saccharomyces cerevisiae*: expression of ethanol tolerance and the *FAD2* gene from *Arabidopsis thaliana*. *Appl. Environ. Microbiol.* 62, 4309–4313.
- Kamada, Y., Jung, U. S., Piotrowski, J., and Levin, D. E. (1995). The protein kinase C-activated MAP kinase pathway of *Saccharomyces cerevisiae* mediates a novel aspect of the heat shock response. *Genes Dev.* 9, 1559–1571. doi:10.1101/gad.9.13.1559
- Keating, J. D., Panganiban, C., and Mansfield, S. D. (2006). Tolerance and adaptation of ethanologenic yeasts to lignocellulosic inhibitory compounds. *Biotechnol. Bioeng.* 93, 1196–1206. doi:10.1002/bit.20838
- Keweloh, H., Weyrauch, G., and Rehm, H.-J. (1990). Phenol-induced membrane changes in free and immobilized *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 33, 66–71. doi:10.1007/bf00170572
- Kim, D., and Hahn, J.-S. (2013). Roles of the Yap1 transcription factor and antioxidants in *Saccharomyces cerevisiae*'s tolerance to furfural and 5-hydroxymethylfurfural, which function as thiol-reactive electrophiles generating oxidative stress. *Appl. Environ. Microbiol.* 79, 5069–5077. doi:10.1128/aem.00643-13
- Kim, I.-S., Kim, Y.-S., Kim, H., Jin, I., and Yoon, H.-S. (2013). *Saccharomyces cerevisiae* KNU5377 stress response during high-temperature ethanol fermentation. *Mol. Cell* 35, 210–218. doi:10.1007/s10059-013-2258-0
- Kim, S.-K., Jin, Y.-S., Choi, I.-G., Park, Y.-C., and Seo, J.-H. (2015). Enhanced tolerance of *Saccharomyces cerevisiae* to multiple lignocellulose-derived inhibitors through modulation of spermidine contents. *Metab. Eng.* 29, 46–55. doi:10.1016/j.ymben.2015.02.004
- Klinke, H. B., Thomsen, A. B., and Ahring, B. K. (2004). Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. *Appl. Microbiol. Biotechnol.* 66, 10–26. doi:10.1007/s00253-004-1642-2
- Kubota, S., Takeo, I., Kume, K., Kanai, M., Shitamukai, A., Mizunuma, M., et al. (2004). Effect of ethanol on cell growth of budding yeast: genes that are important for cell growth in the presence of ethanol. *Biosci. Biotechnol. Biochem.* 68, 968–972. doi:10.1271/bbb.68.968
- Kültz, D. (2003). Evolution of the cellular stress proteome: from monophyletic origin to ubiquitous function. *J. Exp. Biol.* 206, 3119–3124. doi:10.1242/jeb.00549
- Kültz, D. (2005). Molecular and evolutionary basis of the cellular stress response. *Annu. Rev. Physiol.* 67, 225–257. doi:10.1146/annurev.physiol.67.040403.103635
- Kumari, R., and Pramanik, K. (2012). Improvement of multiple stress tolerance in yeast strain by sequential mutagenesis for enhanced bioethanol production. *J. Biosci. Bioeng.* 114, 622–629. doi:10.1016/j.jbiosc.2012.07.007
- Lam, F. H., Ghaderi, A., Fink, G. R., and Stephanopoulos, G. (2014). Engineering alcohol tolerance in yeast. *Science* 346, 71–75. doi:10.1126/science.1257859
- Larsson, S., Palmqvist, E., Hahn-Hägerdal, B., Tengborg, C., Stenberg, K., Zacchi, G., et al. (1999). The generation of fermentation inhibitors during dilute acid hydrolysis of softwood. *Enzyme Microb. Technol.* 24, 151–159. doi:10.1016/S0141-0229(98)00101-X
- Larsson, S., Quintana-Sáinz, A., Reimann, A., Nilvebrant, N.-O., and Jönsson, L. (2000). "Influence of lignocellulose-derived aromatic compounds on oxygen-limited growth and ethanolic fermentation by *Saccharomyces cerevisiae*," in *Twenty-First Symposium on Biotechnology for Fuels and Chemicals*, eds Finkelstein M. and Davison B. (Fort Collins, Colorado: Humana Press), 617–632.
- Lee, K. J., Skotnicki, M. L., Tribe, D. E., and Rogers, P. L. (1980). Kinetic studies on a highly productive strain of *Zymomonas mobilis*. *Biotechnol. Lett.* 2, 339–344. doi:10.1007/bf00138666
- Lee, K.-M., Kalyani, D., Tiwari, M. K., Kim, T.-S., Dhiman, S. S., Lee, J.-K., et al. (2012). Enhanced enzymatic hydrolysis of rice straw by removal of phenolic compounds using a novel laccase from yeast *Yarrowia lipolytica*. *Bioresour. Technol.* 123, 636–645. doi:10.1016/j.biortech.2012.07.066
- Lewis, J. A., Elkon, I. M., Mcgee, M. A., Higbee, A. J., and Gasch, A. P. (2010). Exploiting natural variation in *Saccharomyces cerevisiae* to identify genes for increased ethanol resistance. *Genetics* 186, 1197–1205. doi:10.1534/genetics.110.121871
- Li, B.-Z., and Yuan, Y.-J. (2010). Transcriptome shifts in response to furfural and acetic acid in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 86, 1915–1924. doi:10.1007/s00253-010-2518-2
- Lin, F.-M., Qiao, B., and Yuan, Y.-J. (2009). Comparative proteomic analysis of tolerance and adaptation of ethanologenic *Saccharomyces cerevisiae* to furfural, a lignocellulosic inhibitory compound. *Appl. Environ. Microbiol.* 75, 3765–3776. doi:10.1128/aem.02594-08
- Lindquist, S. (1981). Regulation of protein synthesis during heat shock. *Nature* 293, 311–314. doi:10.1038/293311a0
- Lindquist, S. (1986). The heat-shock response. *Annu. Rev. Biochem.* 55, 1151–1191. doi:10.1146/annurev.bi.55.070186.005443
- Liu, J.-J., Ding, W.-T., Zhang, G.-C., and Wang, J.-Y. (2011). Improving ethanol fermentation performance of *Saccharomyces cerevisiae* in very high-gravity fermentation through chemical mutagenesis and meiotic recombination. *Appl. Microbiol. Biotechnol.* 91, 1239–1246. doi:10.1007/s00253-011-3404-2
- Liu, Z. L., Slininger, P., and Gorsich, S. (2005). "Enhanced biotransformation of furfural and hydroxymethylfurfural by newly developed ethanologenic yeast strains," in *Twenty-Sixth Symposium on Biotechnology for Fuels and Chemicals*, eds Davison B., Evans B., Finkelstein M., and Mcmillan J. (Denver, Colorado: Humana Press), 451–460.
- Lu, Y., Cheng, Y.-F., He, X.-P., Guo, X.-N., and Zhang, B.-R. (2012). Improvement of robustness and ethanol production of ethanologenic *Saccharomyces cerevisiae* under co-stress of heat and inhibitors. *J. Ind. Microbiol. Biotechnol.* 39, 73–80. doi:10.1007/s10295-011-1001-0

- Luong, J. H. T. (1985). Kinetics of ethanol inhibition in alcohol fermentation. *Biotechnol. Bioeng.* 27, 280–285. doi:10.1002/bit.260270311
- Lushchak, V. I. (2011). Adaptive response to oxidative stress: bacteria, fungi, plants and animals. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* 153, 175–190. doi:10.1016/j.cbpc.2010.10.004
- Lv, Y.-J., Wang, X., Ma, Q., Bai, X., Li, B.-Z., Zhang, W., et al. (2014). Proteomic analysis reveals complex metabolic regulation in *Saccharomyces cerevisiae* cells against multiple inhibitors stress. *Appl. Microbiol. Biotechnol.* 98, 2207–2221. doi:10.1007/s00253-014-5519-8
- Lynd, L. R., Cushman, J. H., Nichols, R. J., and Wyman, C. E. (1991). Fuel ethanol from cellulosic biomass. *Science* 251, 1318–1323. doi:10.1126/science.251.4999.1318
- Maris, A. J. A. V., Konings, W. N., Dijken, J. P. V., and Pronk, J. T. (2004). Microbial export of lactic and 3-hydroxypropanoic acid: implications for industrial fermentation processes. *Metab. Eng.* 6, 245–255. doi:10.1016/j.ymben.2004.05.001
- Meena, R., Thakur, S., and Chakrabarti, A. (2011). Regulation of *Saccharomyces cerevisiae* plasma membrane H<sup>+</sup>-ATPase (Pma1) by dextrose and Hsp30 during exposure to thermal stress. *Indian J. Microbiol.* 51, 153–158. doi:10.1007/s12088-011-0137-y
- Miller, E. N., Jarboe, L. R., Yomano, L. P., York, S. W., Shanmugam, K. T., and Ingram, L. O. (2009). Silencing of NADPH-dependent oxidoreductase genes (*yqhD* and *dkgA*) in furfural-resistant ethanologenic *Escherichia coli*. *Appl. Environ. Microbiol.* 75, 4315–4323. doi:10.1128/aem.00567-09
- Mills, C., Child, J. J., and Spencer, J. F. T. (1971). The utilization of aromatic compounds by yeasts. *Antonie Van Leeuwenhoek* 37, 281–287. doi:10.1007/bf02218497
- Mishra, P., and Prasad, R. (1989). Relationship between ethanol tolerance and fatty acyl composition of *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 30, 294–298. doi:10.1007/bf00256221
- Mitchell, A., Romano, G. H., Groisman, B., Yona, A., Dekel, E., Kupiec, M., et al. (2009). Adaptive prediction of environmental changes by microorganisms-supplementary info available. *Nature* 460, 220–224. doi:10.1038/nature08112
- Nagodawithana, T. W., and Steinkraus, K. H. (1976). Influence of the rate of ethanol production and accumulation on the viability of *Saccharomyces cerevisiae* in “rapid fermentation”. *Appl. Environ. Microbiol.* 31, 158–162.
- Nakanishi, A., Bae, J., Fukai, K., Kukumoto, N., Kuroda, K., Ogawa, J., et al. (2012). Effect of pretreatment of hydrothermally processed rice straw with laccase-displaying yeast on ethanol fermentation. *Appl. Microbiol. Biotechnol.* 94, 939–948. doi:10.1007/s00253-012-3876-8
- Narendranath, N. V., and Power, R. (2005). Relationship between pH and medium dissolved solids in terms of growth and metabolism of *Lactobacilli* and *Saccharomyces cerevisiae* during ethanol production. *Appl. Environ. Microbiol.* 71, 2239–2243. doi:10.1128/aem.71.5.2239-2243.2005
- Neves, M. J., and Francois, J. (1992). On the mechanism by which a heat shock induces trehalose accumulation in *Saccharomyces cerevisiae*. *Biochem. J.* 288, 859–864. doi:10.1042/bj2880859
- Nwaka, S., Kopp, M., Burgert, M., Deuchler, I., Kienle, I., and Holzer, H. (1994). Is thermotolerance of yeast dependent on trehalose accumulation? *FEBS Lett.* 344, 225–228. doi:10.1016/0014-5793(94)00385-8
- Nwaka, S., Mechler, B., Destruelle, M., and Holzer, H. (1995). Phenotypic features of trehalase mutants in *Saccharomyces cerevisiae*. *FEBS Lett.* 360, 286–290. doi:10.1016/0014-5793(95)00105-1
- Olsson, L., and Hahn-Hägerdal, B. (1993). Fermentative performance of bacteria and yeasts in lignocellulose hydrolysates. *Proc. Biochem.* 28, 249–257. doi:10.1016/0032-9592(93)80041-E
- Palmqvist, E., Almeida, J. S., and Hahn-Hägerdal, B. (1999a). Influence of furfural on anaerobic glycolytic kinetics of *Saccharomyces cerevisiae* in batch culture. *Biotechnol. Bioeng.* 62, 447–454. doi:10.1002/(SICI)1097-0290(19990220)62:4<447::AID-BIT7>3.0.CO;2-0
- Palmqvist, E., Grage, H., Meinander, N. Q., and Hahn-Hägerdal, B. (1999b). Main and interaction effects of acetic acid, furfural, and p-hydroxybenzoic acid on growth and ethanol productivity of yeasts. *Biotechnol. Bioeng.* 63, 46–55. doi:10.1002/(SICI)1097-0290(19990405)63:1<46::AID-BIT5>3.3.CO;2-A
- Palmqvist, E., and Hahn-Hägerdal, B. (2000a). Fermentation of lignocellulosic hydrolysates. I: inhibition and detoxification. *Bioresour. Technol.* 74, 17–24. doi:10.1016/S0960-8524(99)00160-1
- Palmqvist, E., and Hahn-Hägerdal, B. (2000b). Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. *Bioresour. Technol.* 74, 25–33. doi:10.1016/S0960-8524(99)00161-3
- Pampulha, M. E., and Loureiro-Dias, M. C. (1989). Combined effect of acetic acid, pH and ethanol on intracellular pH of fermenting yeast. *Appl. Microbiol. Biotechnol.* 31, 547–550. doi:10.1007/bf00270792
- Pampulha, M. E., and Loureiro-Dias, M. C. (1990). Activity of glycolytic enzymes of *Saccharomyces cerevisiae* in the presence of acetic acid. *Appl. Microbiol. Biotechnol.* 34, 375–380. doi:10.1007/bf00170063
- Pan, X., Gilkes, N., Kadla, J., Pye, K., Saka, S., Gregg, D., et al. (2006). Bioconversion of hybrid poplar to ethanol and co-products using an organosolv fractionation process: optimization of process yields. *Biotechnol. Bioeng.* 94, 851–861. doi:10.1002/bit.20905
- Panchal, C. J., and Stewart, G. G. (1980). The effect of osmotic pressure on the production and excretion of ethanol and glycerol by a brewing yeast strain. *J. Inst. Brew.* 86, 207–210. doi:10.1002/j.2050-0416.1980.tb06867.x
- Petersson, A., Almeida, J. R. M., Modig, T., Karhumaa, K., Hahn-Hägerdal, B., Gorwa-Grauslund, M. F., et al. (2006). A 5-hydroxymethyl furfural reducing enzyme encoded by the *Saccharomyces cerevisiae* ADH6 gene conveys HMF tolerance. *Yeast* 23, 455–464. doi:10.1002/yea.1370
- Pinel, D., Colatriano, D., Jiang, H., Lee, H., and Martin, V. (2015). Deconstructing the genetic basis of spent sulphite liquor tolerance using deep sequencing of genome-shuffled yeast. *Biotechnol. Biofuels* 8, 53. doi:10.1186/s13068-015-0241-z
- Pinel, D., D'aoust, E., Del Cardayre, S. B., Bajwa, P. K., Lee, H., and Martin, V. J. J. (2011). *Saccharomyces cerevisiae* genome shuffling through recursive population mating leads to improved tolerance to spent sulfite liquor. *Appl. Environ. Microbiol.* 77, 4736–4743. doi:10.1128/aem.02769-10
- Piper, P. W. (1993). Molecular events associated with acquisition of heat tolerance by the yeast *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* 11, 339–355. doi:10.1111/j.1574-6976.1993.tb00005.x
- Piper, P. W., Ortiz-Calderon, C., Holyoak, C., Coote, P., and Cole, M. (1997). Hsp30, the integral plasma membrane heat shock protein of *Saccharomyces cerevisiae*, is a stress-inducible regulator of plasma membrane H<sup>+</sup>-ATPase. *Cell Stress Chaperones* 2, 12–24. doi:10.1379/1466-1268(1997)002<0012:HTIPMH>2.3.CO;2
- Postmus, J., Aardema, R., De Koning, L. J., De Koster, C. G., Brul, S., and Smits, G. J. (2012). Isoenzyme expression changes in response to high temperature determine the metabolic regulation of increased glycolytic flux in yeast. *FEMS Yeast Res.* 12, 571–581. doi:10.1111/j.1567-1364.2012.00807.x
- Postmus, J., Canelas, A. B., Bouwman, J., Bakker, B. M., Van Gulik, W., De Mattos, M. J. T., et al. (2008). Quantitative analysis of the high temperature-induced glycolytic flux increase in *Saccharomyces cerevisiae* reveals dominant metabolic regulation. *J. Biol. Chem.* 283, 23524–23532. doi:10.1074/jbc.M802908200
- Richard, P., Viljanen, K., and Penttilä, M. (2015). Overexpression of PAD1 and FDC1 results in significant cinnamic acid decarboxylase activity in *Saccharomyces cerevisiae*. *AMB Express* 5, 12. doi:10.1186/s13568-015-0103-x
- Rodicio, R., and Heinisch, J. J. (2010). Together we are strong – cell wall integrity sensors in yeasts. *Yeast* 27, 531–540. doi:10.1002/yea.1785
- Roosen, J., Engelen, K., Marchal, K., Mathys, J., Griffioen, G., Camerini, E., et al. (2005). PKA and Sch9 control a molecular switch important for the proper adaptation to nutrient availability. *Mol. Microbiol.* 55, 862–880. doi:10.1111/j.1365-2958.2004.04429.x
- Ruis, H., and Schüller, C. (1995). Stress signaling in yeast. *Bioessays* 17, 959–965. doi:10.1002/bies.950171109
- Sá-Correia, I., and Van Uden, N. (1983). Temperature profiles of ethanol tolerance: effects of ethanol on the minimum and the maximum temperatures for growth of the yeasts *Saccharomyces cerevisiae* and *Kluyveromyces fragilis*. *Biotechnol. Bioeng.* 25, 1665–1667. doi:10.1002/bit.260250620
- Sanchez, Y., Taulien, J., Borkovich, K. A., and Lindquist, S. (1992). *Hsp104* is required for tolerance to many forms of stress. *EMBO J.* 11, 2357–2364.
- Sárvári Horváth, I., Franzén, C. J., Taherzadeh, M. J., Niklasson, C., and Lidén, G. (2003). Effects of furfural on the respiratory metabolism of *Saccharomyces cerevisiae* in glucose-limited chemostats. *Appl. Environ. Microbiol.* 69, 4076–4086. doi:10.1128/aem.69.7.4076-4086.2003
- Sasano, Y., Watanabe, D., Ukibe, K., Inai, T., Ohtsu, I., Shimoi, H., et al. (2012). Overexpression of the yeast transcription activator Msn2 confers furfural resistance and increases the initial fermentation rate in ethanol production. *J. Biosci. Bioeng.* 113, 451–455. doi:10.1016/j.jbiosc.2011.11.017
- Scrimale, T., Didone, L., De Mesy Bentley, K. L., and Krysan, D. J. (2009). The unfolded protein response is induced by the cell wall integrity mitogen-activated protein kinase signaling cascade and is required for cell wall integrity



- in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 20, 164–175. doi:10.1091/mbc.E08-08-0809
- Shen, Y., Li, H., Wang, X., Zhang, X., Hou, J., Wang, L., et al. (2014). High vanillin tolerance of an evolved *Saccharomyces cerevisiae* strain owing to its enhanced vanillin reduction and antioxidative capacity. *J. Ind. Microbiol. Biotechnol.* 41, 1637–1645. doi:10.1007/s10295-014-1515-3
- Sehnem, N. T., Da Silva Machado, A., Leite, F. C. B., De Barros Pita, W., De Morais, M. A. Jr., and Ayub, M. A. Z. (2013). 5-hydroxymethylfurfural induces ADH7 and AR11 expression in tolerant industrial *Saccharomyces cerevisiae* strain P6H9 during bioethanol production. *Bioresour. Technol.* 133, 190–196. doi:10.1016/j.biortech.2013.01.063
- Shama, S., Lai, C.-Y., Antoniazzi, J. M., Jiang, J. C., and Jazwinski, S. M. (1998). Heat stress-induced life span extension in yeast. *Exp. Cell Res.* 245, 379–388. doi:10.1006/excr.1998.4279
- Shi, D.-J., Wang, C.-L., and Wang, K.-M. (2009). Genome shuffling to improve thermotolerance, ethanol tolerance and ethanol productivity of *Saccharomyces cerevisiae*. *J. Ind. Microbiol. Biotechnol.* 36, 139–147. doi:10.1007/s10295-008-0481-z
- Shi, L., and Tu, B. P. (2013). Acetyl-CoA induces transcription of the key G1 cyclin CLN3 to promote entry into the cell division cycle in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 7318–7323. doi:10.1073/pnas.1302490110
- Shin, D. Y., Matsumoto, K., Iida, H., Uno, I., and Ishikawa, T. (1987). Heat shock response of *Saccharomyces cerevisiae* mutants altered in cyclic AMP-dependent protein phosphorylation. *Mol. Cell Biol.* 7, 244–250. doi:10.1128/mcb.7.1.244
- Singer, M. A., and Lindquist, S. (1998). Thermotolerance in *Saccharomyces cerevisiae*: the Yin and Yang of trehalose. *Trends Biotechnol.* 16, 460–468. doi:10.1016/s0167-7799(98)01251-7
- Smith, J., Van Rensburg, E., and Gorgens, J. (2014). Simultaneously improving xylose fermentation and tolerance to lignocellulosic inhibitors through evolutionary engineering of recombinant *Saccharomyces cerevisiae* harbouring xylose isomerase. *BMC Biotechnol.* 14:41. doi:10.1186/1472-6750-14-41
- Stemmer, W. P. (1994). DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. U.S.A.* 91, 10747–10751. doi:10.1073/pnas.91.22.10747
- Suutari, M., Liukkonen, K., and Laakso, S. (1990). Temperature adaptation in yeasts: the role of fatty acids. *J. Gen. Microbiol.* 136, 1469–1474. doi:10.1099/00221287-136-8-1469
- Suutari, M., Rintamäki, A., and Laakso, S. (1997). Membrane phospholipids in temperature adaptation of *Candida utilis*: alterations in fatty acid chain length and unsaturation. *J. Lipid Res.* 38, 790–794.
- Tagkopoulos, I., Liu, Y.-C., and Tavazoie, S. (2008). Predictive behavior within microbial genetic networks. *Science* 320, 1313–1317. doi:10.1126/science.1154456
- Taherzadeh, M. J., Gustafsson, L., Niklasson, C., and Lidén, G. (1999). Conversion of furfural in aerobic and anaerobic batch fermentation of glucose by *Saccharomyces cerevisiae*. *J. Biosci. Bioeng.* 87, 169–174. doi:10.1016/S1389-1723(99)89007-0
- Takahara, T., and Maeda, T. (2012). Transient sequestration of TORC1 into stress granules during heat stress. *Mol. Cell* 47, 242–252. doi:10.1016/j.molcel.2012.05.019
- Tao, X., Zheng, D., Liu, T., Wang, P., Zhao, W., Zhu, M., et al. (2012). A novel strategy to construct yeast *Saccharomyces cerevisiae* strains for very high gravity fermentation. *PLoS ONE* 7:e31235. doi:10.1371/journal.pone.0031235
- Temple, M. D., Perrone, G. G., and Dawes, I. W. (2005). Complex cellular responses to reactive oxygen species. *Trends Cell Biol.* 15, 319–326. doi:10.1016/j.tcb.2005.04.003
- Tokiwa, G., Tyers, M., Volpe, T., and Futcher, B. (1994). Inhibition of G1 cyclin activity by the Ras/cAMP pathway in yeast. *Nature* 371, 342–345. doi:10.1038/371342a0
- Urban, J., Soulard, A., Huber, A., Lippman, S., Mukhopadhyay, D., Deloche, O., et al. (2007). Sch9 is a major target of TORC1 in *Saccharomyces cerevisiae*. *Mol. Cell* 26, 663–674. doi:10.1016/j.molcel.2007.04.020
- Verduyn, C., Postma, E., Scheffers, W. A., and Van Dijken, J. P. (1992). Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* 8, 501–517. doi:10.1002/yea.320080703
- Verna, J., Lodder, A., Lee, K., Vagts, A., and Ballester, R. (1997). A family of genes required for maintenance of cell wall integrity and for the stress response in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 94, 13804–13809. doi:10.1073/pnas.94.25.13804
- Viegas, C. A., and Sá-Correia, I. (1991). Activation of plasma membrane ATPase of *Saccharomyces cerevisiae* by octanoic acid. *J. Gen. Microbiol.* 137, 645–651. doi:10.1099/00221287-137-3-645
- Vilella, F., Herrero, E., Torres, J., and De La Torre-Ruiz, M. A. (2005). Pkc1 and the upstream elements of the cell integrity pathway in *Saccharomyces cerevisiae*, Rom2 and Mtl1, are required for cellular responses to oxidative stress. *J. Biol. Chem.* 280, 9149–9159. doi:10.1074/jbc.M411062200
- Wallace-Salinas, V., and Gorwa-Grauslund, M. (2013). Adaptive evolution of an industrial strain of *Saccharomyces cerevisiae* for combined tolerance to inhibitors and temperature. *Biotechnol. Biofuels* 6, 151. doi:10.1186/1754-6834-6-151
- Wojda, I., Alonso-Monge, R., Bebelman, J.-P., Mager, W. H., and Siderius, M. (2003). Response to high osmotic conditions and elevated temperature in *Saccharomyces cerevisiae* is controlled by intracellular glycerol and involves coordinate activity of MAP kinase pathways. *Microbiology* 149, 1193–1204. doi:10.1099/mic.0.26110-0
- Yamamoto, A., Ueda, J., Yamamoto, N., Hashikawa, N., and Sakurai, H. (2007). Role of heat shock transcription factor in *Saccharomyces cerevisiae* oxidative stress response. *Eukaryot. Cell* 6, 1373–1379. doi:10.1128/EC.00098-07
- Yomano, L. P., York, S. W., and Ingram, L. O. (1998). Isolation and characterization of ethanol-tolerant mutants of *Escherichia coli* KO11 for fuel ethanol production. *J. Ind. Microbiol. Biotechnol.* 20, 132–138. doi:10.1038/sj.jim.2900496
- Yona, A. H., Manor, Y. S., Herbst, R. H., Romano, G. H., Mitchell, A., Kupiec, M., et al. (2012). Chromosomal duplication is a transient evolutionary solution to stress. *Proc. Natl. Acad. Sci. U.S.A.* 109, 21010–21015. doi:10.1073/pnas.1211150109
- Zaldivar, J., Martinez, A., and Ingram, L. O. (1999). Effect of selected aldehydes on the growth and fermentation of ethanologenic *Escherichia coli*. *Biotechnol. Bioeng.* 65, 24–33. doi:10.1002/(SICI)1097-0290(19991005)65:1<24::AID-BIT4>3.0.CO;2-2
- Zhang, Y.-X., Perry, K., Vinci, V. A., Powell, K., Stemmer, W. P. C., and Del Cardayre, S. B. (2002). Genome shuffling leads to rapid phenotypic improvement in bacteria. *Nature* 415, 644–646. doi:10.1038/415644a

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Caspeta, Castillo and Nielsen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Systems Biology Approaches to Understand Natural Products Biosynthesis

Cuauhtemoc Licona-Cassani<sup>1,2</sup>, Pablo Cruz-Morales<sup>2</sup>, Angel Manteca<sup>3</sup>, Francisco Barona-Gomez<sup>2</sup>, Lars K. Nielsen<sup>1</sup> and Esteban Marcellin<sup>1\*</sup>

<sup>1</sup> Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland, Brisbane, QLD, Australia, <sup>2</sup> National Laboratory of Genomics for Biodiversity (LANGEBIO), Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (Cinvestav-IPN), Irapuato, México, <sup>3</sup> Departamento de Biología Funcional and Instituto Universitario de Oncología del Principado de Asturias (IUOPA), Facultad de Medicina, Universidad de Oviedo, Oviedo, Spain

## OPEN ACCESS

### Edited by:

Alvaro R. Lara,  
Universidad Autónoma  
Metropolitana-Cuajimalpa, Mexico

### Reviewed by:

Preetam Ghosh,  
Virginia Commonwealth University,  
USA  
Yaojun Tong,  
Chinese Academy of Sciences, China  
Irina Borodina,  
Technical University of Denmark,  
Denmark

### \*Correspondence:

Esteban Marcellin  
e.marcellin@uq.edu.au

### Specialty section:

This article was submitted to  
Systems Biology,  
a section of the journal  
Frontiers in Bioengineering and  
Biotechnology

Received: 25 July 2015

Accepted: 24 November 2015

Published: 09 December 2015

### Citation:

Licona-Cassani C, Cruz-Morales P,  
Manteca A, Barona-Gomez F,  
Nielsen LK and Marcellin E (2015)  
Systems Biology Approaches to  
Understand Natural Products  
Biosynthesis.  
Front. Bioeng. Biotechnol. 3:199.  
doi: 10.3389/fbioe.2015.00199

Actinomycetes populate soils and aquatic sediments that impose biotic and abiotic challenges for their survival. As a result, actinomycetes metabolism and genomes have evolved to produce an overwhelming diversity of specialized molecules. Polyketides, non-ribosomal peptides, post-translationally modified peptides, lactams, and terpenes are well-known bioactive natural products with enormous industrial potential. Accessing such biological diversity has proven difficult due to the complex regulation of cellular metabolism in actinomycetes and to the sparse knowledge of their physiology. The past decade, however, has seen the development of *omics* technologies that have significantly contributed to our better understanding of their biology. Key observations have contributed toward a shift in the exploitation of actinomycete's biology, such as using their full genomic potential, activating entire pathways through key metabolic elicitors and pathway engineering to improve biosynthesis. Here, we review recent efforts devoted to achieving enhanced discovery, activation, and manipulation of natural product biosynthetic pathways in model actinomycetes using genome-scale biological datasets.

**Keywords:** actinomycetes, genome mining, genomics, transcriptomics, proteomics, metabolomics, genome-scale metabolic reconstructions

## INTRODUCTION

Actinomycetes represent one of the largest bacterial phyla and are primary contributors to carbon cycling and a major source of bioactive natural products (BNP) including, most prominently, antibiotics. Despite their prime importance, our understanding of actinomycete's biology remains elusive owing to a characteristically large, convoluted, high GC content genome (Demain, 2014). The complexity of actinomycetes genomes was only fully revealed in the last decade as part of the genomic revolution. Sequencing of the first actinomycetes genomes revealed a plethora of bioactive secondary metabolites yet to be discovered in addition to the well-characterized biosynthetic gene clusters (BGC) (Doroghazi et al., 2014). According to NCBI database, to date around 1,000 actinomycete genomes have been fully sequenced and annotated. Homology sequence-based bioinformatic tools have confirmed their great potential as BNP producers; for example, species

of *Streptomyces*, *Salinispora*, and *Saccharopolyspora* families contain an average of 30 secondary metabolite gene clusters (Nett et al., 2009).

The physiological changes leading to BNP biosynthesis in actinomycetes have been thoroughly studied over the past 10 years. Considerable work has advanced our understanding of the transitional stage triggering BNP biosynthesis (also known as the “metabolic switch”; Alam et al., 2010) and with it, our understanding of the physiological changes leading to secondary pathways activation. However, a lack of full understanding of this physiological transition stage has prevented us from manipulating fully this cellular process using metabolic engineering. Here, we review landmark studies contributing to the discovery, activation, and manipulation of metabolic pathways for BNP through the development of genome-wide biological datasets and systems biology in actinomycetes (Figure 1).

## PATHWAY DISCOVERY: FROM IMPROVED GENOME ANNOTATION TO THE DISCOVERY OF NEW BNP BIOSYNTHETIC PATHWAYS

Genome annotation is the basis for pathway discovery and manipulation. Pathway discovery typically follows a defined pipeline: genome sequencing, annotation, gene discovery, and pathway manipulation. The exponential increase in sequencing efficiency is yielding an ever-increasing number of sequenced genomes, causing a bottleneck due to an often limited understanding of genome sequences. *In silico* approaches mainly rely on sequence homology scores to experimentally characterized sequences. Historically, however, functional microbiology has focused on a handful of microorganisms. Therefore, the genomic space for *in silico* genome annotation pipelines is biased for certain G + C content sequences, gene length, and organization. For instance, approximately 60% of the bacterial genomic space is miss-annotated in terms of gene boundaries (start/stop codons) caused by minimal cross-checks between computationally assigned open-reading frames (ORFs) and real genes (Nielsen and Krogh, 2005).

Bioinformatics-based pipelines failed to annotate accurately short-length proteins and high G + C content sequences in an annotation effort for 46 bacterial and archaea genomes (Venter et al., 2011). By contrast, functional annotations supported by “omics technologies” dramatically improve gene function assignment, particularly in less characterized microorganisms such as *Geobacter sulfurreducens* (Qiu et al., 2010) or the erythromycin producer actinomycete *Saccharopolyspora erythraea* (Marcellin et al., 2013a). Integration of proteomics and transcriptomics approaches has led to the re-annotation of these genomes, allowing for correction of hundreds of gene boundaries, the confirmation of hypothetical proteins and the discovery of dozens of new genes. A combination of proteomics and genomics, also known as proteogenomics, has also been

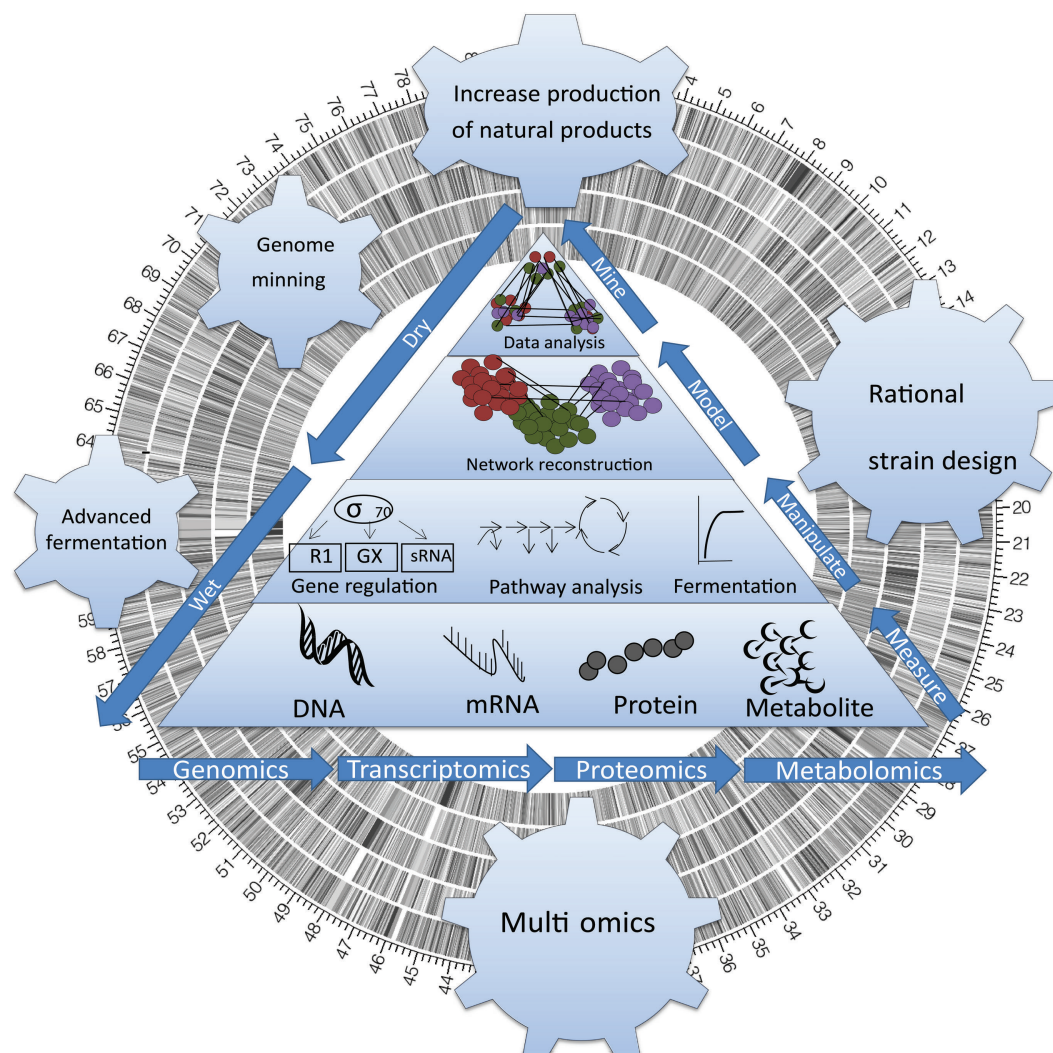
used to deliver unbiased correlations between genome sequence and protein expression (Gupta et al., 2007; Gallien et al., 2009; Armengaud, 2010; Castellana and Bafna, 2010; Marcellin et al., 2013a).

## Genome Mining and Pathway Discovery

The first sequenced genomes of BNP producers were *Streptomyces coelicolor* (Bentley et al., 2002), *Streptomyces avermitilis* (Ikeda et al., 2003), producer of the insecticide/anthelmintic avermectin, *S. erythraea* (Oliynyk et al., 2007) and producer of the classic antibiotic streptomycin *Streptomyces griseus* (Ohnishi et al., 2008). Further sequence inspection of such genomes and other model actinomycetes have opened a plethora of BGCs, and revealed the great potential of actinomycetes genomes for the production of BNPs (Nett et al., 2009; Aigle et al., 2014; Doroghazi et al., 2014; Ikeda et al., 2014). However, exploiting this rich source of BNP has proven challenging. Genomic analyses show an abundance of known BGCs (i.e., chemically and genetically known), hypothetical BCGs (i.e., chemically unknown – genetically known), and cryptic BCGs (i.e., chemically unknown – genetically unknown) (Zerikly and Challis, 2009; Doroghazi and Metcalf, 2013).

Initial approaches to the discovery and identification of BNP were based on the search for cryptic BGC. The most common method involves gene mapping of enzymatic assembly complexes such as polyketide synthases (PKSs), non-ribosomal peptide synthases (NRPSs), and other enzymes typically related to BNPs (e.g., lanthipeptide synthases, terpene synthases, etc.). While simplistic, accumulation of structural, mechanistic, genetic, and chemical information on PKs and NRPs has allowed for the prediction of structures and chemical properties of dozens of BCGs from DNA sequences (Walsh et al., 2006; Hertweck, 2009; Jenke-Kodama and Dittmann, 2009; Koglin and Walsh, 2009; Walsh and Fischbach, 2010). Incorporating these mining strategies in specialized bioinformatic pipelines has revolutionized the genome mining scene efficiency. Genome-scale prediction of putative BGCs is nowadays possible within a few of hours.

Continuous progress has enabled the emergence of bioinformatics platforms, such as CLUSEAN (Weber et al., 2009), ClustScan (Starcevic et al., 2008), np.searcher (Li et al., 2009), SMURF (Khaldi et al., 2010), and antiSMASH (Medema et al., 2011b; Blin et al., 2013; Weber et al., 2015). The latter is the most popular system for automated BNP genome mining since it analyzes BGC domains to propose *loci*, chemical scaffold, and putative chemical structures. However, one of the biggest disadvantages of the use of these genome mining approaches is their intrinsic limitations to BGCs from known chemical structures. Complementary approaches have emerged to enable the discovery of novel BCGs, such as ClusterFinder (Cimermanic et al., 2014) or EvoMining (Medema and Fischbach, 2015). ClusterFinder uses hidden Markov model-based algorithms and Pfam as search database to annotate BCGs by clusters of protein domains with a biosynthetic logic. The use of ClusterFinder has allowed the detection of previously unknown classes of BCGs (Cimermanic et al., 2014). On the other hand, EvoMining is



**FIGURE 1 | Sequencing actinomycete genomes has revealed an unexpected complexity.** Systems Biology has opened access to the untapped chemical diversity encoded within the global microbial genome, including the vast majority (>99%) of taxa that are currently deemed unculturable, and a wealth of bioactive genes that are currently silent (untranslated) under standard cultivation condition.

a functional phylogenomic pipeline that identifies expanded, repurposed enzyme families, with the potential to catalyze new conversions within BGC (Medema and Fischbach, 2015). This innovative method embraces the predictive power of evolutionary theory leading to model-independent predictions that include gene clusters that do not follow traditional biosynthetic rules. The method has been used for the discovery of the genes directing synthesis of small peptide aldehydes and the first bio-synthetic system for arseno-organic metabolites.

Overall, genomic approaches have significantly improved the prediction of BNP from unannotated sequences and provided deep insights into the identification of novel chemical species. The genomic approach is limited to the known repertoire of BCGs, ignoring regulatory information for pathway activation.

## PATHWAY ACTIVATION: SYSTEMS BIOLOGY ANALYSIS OF ACTINOMYCETE PHYSIOLOGY AND DEVELOPMENT

Systems biology protocols have been successfully used to describe germination (Piette et al., 2005; Yagüe et al., 2013a; Bobek et al., 2014), programmed cell death (Manteca et al., 2005), diauxic lag phase (Novotna et al., 2003), mutant analyses (bald A mutant) (Kim et al., 2005; Hesketh et al., 2007), and phosphate limitation (Rodríguez-García et al., 2007). Given that biosynthesis of natural products in actinomycetes is conceived as a physiological response to environmental changes (e.g., change of temperature, nutritional conditions, etc.), it is assumed that understanding their physiological behavior would provide the lead for natural product pathway activation and manipulation. Here, we focus on



reviewing efforts devoted to understand the physiological transitions prior the activation of known natural product biosynthesis and the approaches used for the activation of unknown natural products biosynthetic pathways in model actinomycetes.

## Physiological Transitions and Development

Actinomycetes undergo drastic physiological changes during their developmental cycle (i.e., programed cell death and sporulation). In contrast to previous assumptions that sporulation events exclusively occurred in solid cultures (Flardh and Buttner, 2009), differentiation during pre-sporulation stages have been described in both solid and liquid *Streptomyces* cultures (Manteca et al., 2010). The existence of two different mycelia (MI and MII) across the developmental cycle has been characterized using iTRAQ LC-MS/MS proteomics, phosphoproteomics, and microarray-based transcriptomics (Manteca et al., 2010, 2011; Yagüe et al., 2013b). Specifically, proteins involved in antibiotic biosynthesis were upregulated in MII, and primary metabolism proteins from glycolysis, protein biosynthesis, and tricarboxylic acid cycle were upregulated in the MI. The second multinucleated mycelium with (aerial) and without (substrate) hydrophobic covers constituted a unique reproductive structure (Manteca et al., 2010). The most remarkable differences between MII from solid and liquid cultures involved proteins regulating the final stages of hyphae compartmentalization and spore formation (Manteca et al., 2010).

Similarly, characterization of the *S. erythraea* developmental cycle in bioreactors has been explored at base resolution transcription (RNA-seq), proteome (iTRAQ) and phosphoproteome (sMRM) (Marcellin et al., 2013a,b; Licona-Cassani et al., 2014) (Figure 2). The studies focused on the metabolic switch, a distinct transformational event that bisects two growth phases in actinomycetes and is characterized by rapid molecular and morphological changes. Authors found that the *S. erythraea* transcriptome undergoes extensive events of targeted mRNA degradation and transcription of mRNAs for adaptive metabolic functions, thereby resetting cells for the induction of a replacement transcriptional program. A suite of RNase and proteases mediate a targeted destruction of the transcriptome and proteome (suicidal patterns) in concert with the shifting of broad transcription macro-domains, delineated by core/non-core genomic regions. In addition, the temporal-dynamic, semi-quantitative phosphoproteomic study revealed that proteins from central metabolism (putative acetyl-CoA carboxylase, isocitrate lyase, and 2-oxoglutarate dehydrogenase) and key developmental pathways (trypsin-like serine protease, ribonuclease Rne/Rng, and ribosomal proteins) in *S. erythraea* change dramatically the degree of phosphorylation across the developmental cycle in liquid cultures (Figure 2) (Licona-Cassani et al., 2014).

One of the most significant observations linking actinomycete physiological behavior and pathway activation was made by Nieselt and collaborators (Nieselt et al., 2010). Using a temporal-dynamic transcriptomic analysis, Nieselt and collaborators identified the existence of several transitional stages along the fermentation that coincide with activation of natural product metabolic pathways

in *S. coelicolor* (Nieselt et al., 2010). Under their bioreactor settings, early coordinated gene expression changes of genes related to nitrogen metabolism, including glutamine synthases I and II and the signaling protein GlnK is observed under similar temporal space as genes from the CPK antibiotic biosynthetic pathway. Interestingly, such transcriptional changes were observed under nitrogen sufficiency conditions. In addition, an unexpected transcriptional switch for developmental genes, such as chaplins, bldN, and whiH was registered showing for the first time that developmental genes are transcribed in *S. coelicolor* liquid cultures. Finally, the traditional metabolic switch was observed by a strong upregulation of the *pho* regulon together with the upregulation of the pigmented antibiotic undecylprodigiosin and actinorhodin (Nieselt et al., 2010).

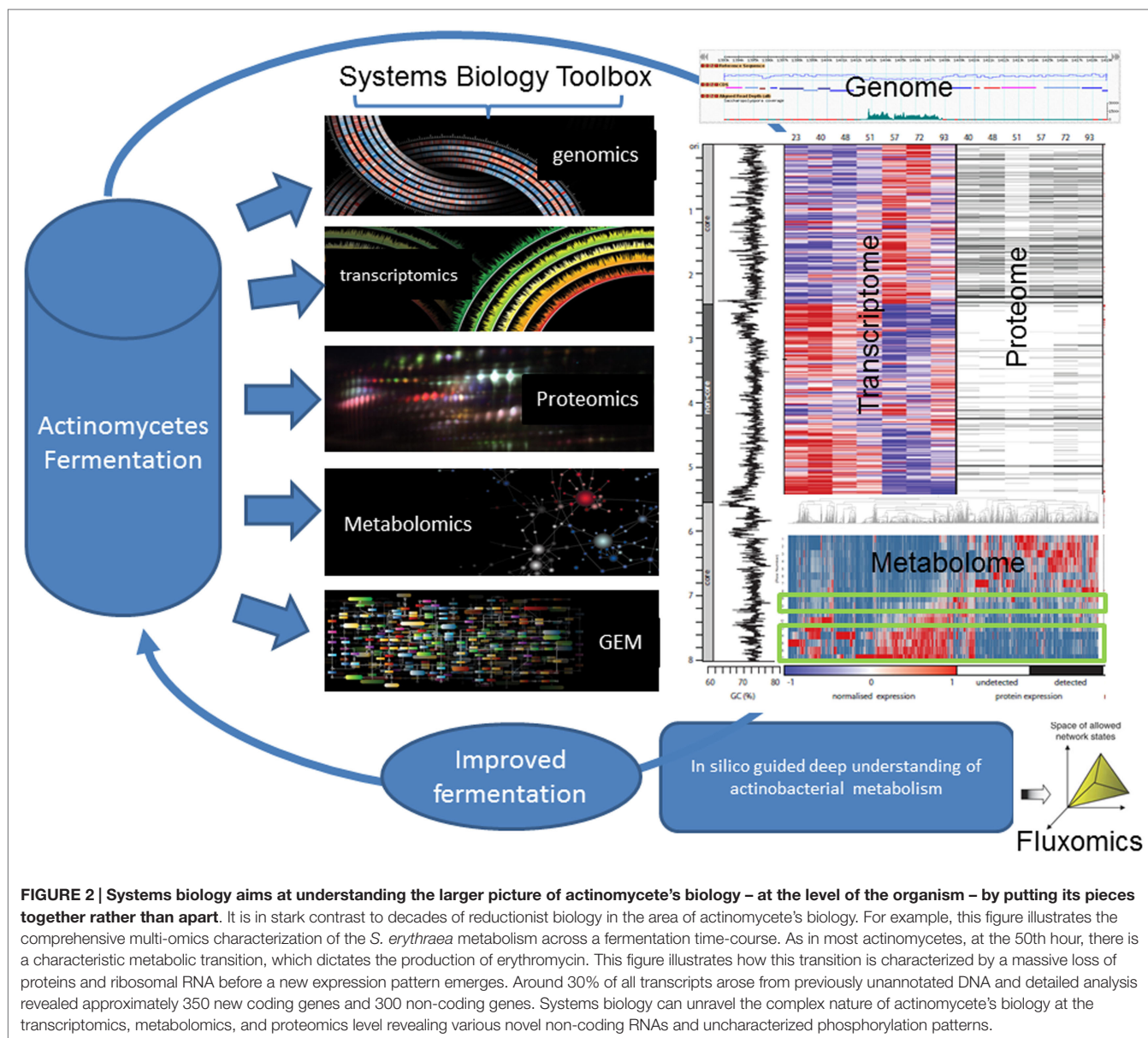
While we are still far from overcoming the physiological barrier of achieving pathway activation and exploiting the full genomic potential of actinomycetes, systems biology approaches have significantly contributed to shifting key paradigms. First, we know that pathway activation does not follow the same regulatory rules to model actinomycetes (e.g., *S. coelicolor* or *Mycobacterium tuberculosis*); in fact, it is now possible to understand such differences in a single experiment. More importantly, systems biology has exposed a subset of strain-metabolite-specific regulatory mechanisms such as non-coding RNAs (Marcellin et al., 2013b), dynamic phosphorylation of ribosomal proteins (Licona-Cassani et al., 2014), acetylation of RNA degradation-related proteins (Huang et al., 2015), among others. The last part of this mini-review focuses on the efforts to manipulate metabolic pathways in actinomycetes using genome-scale metabolic reconstructions and metabolic engineering strategies.

## PATHWAY MANIPULATION: FROM RATIONAL DESIGN TO GENOME-SCALE MODEL – GUIDED METABOLIC ENGINEERING STRATEGIES IN ACTINOMYCETES

Production processes with sub-hundreds of mg/L product titers are considered unsustainable for industrial scale production. Actinomycetes cultures are also slow growing and unpredictable even under controlled conditions (i.e., bioreactor fermentations), and as such are difficult to ferment. For over 50 years, entire teams of metabolic and bioprocess engineers have used classical approaches such as random mutagenesis rounds (Tanaka et al., 2009; Jung et al., 2011), media design, and process optimization (Hamedi et al., 2004; El-Enshasy et al., 2008; Zou et al., 2009), and rationally designed metabolic engineering strategies (Reeves et al., 2006, 2007; Olano et al., 2008) to engineer actinomycetes bioprocesses and strains to achieve acceptable titers. Modern strain engineering uses genome-scale models (GEMs) in combination with *omics* data for the integration of genome-scale biological datasets toward the manipulation of metabolism.

Since the first GEM release, more than a decade ago (Edwards and Palsson, 1999), applications have expanded from *in silico* metabolic predictions (Edwards and Palsson, 2000; Schilling et al., 2002; Park et al., 2007) to the discovery of antibacterial





targets (Oberhardt et al., 2009; Kim et al., 2010), integration of biological datasets (Lerman et al., 2012; Imam et al., 2015), phenotype prediction (O'Brien et al., 2013), estimation of metabolic capabilities (O'Brien et al., 2015), and the study of evolutionary relationships of metabolic and regulatory networks (Oberhardt et al., 2009; Kim et al., 2010; Barona-Gomez et al., 2012).

While draft GEM are now routinely generated using high-throughput automated pipelines (Aziz et al., 2008; Henry et al., 2010), successful applications of GEMs to microbial metabolism only have occurred on exhaustively (manually) curated/experimentally validated metabolic reconstructions. In this regard, despite being our primary microbial source for antibiotics, only a handful of manually curated GEM for actinomycetes have been reported (Borodina et al., 2005a; Beste et al., 2007; Jamshidi and

Palsson, 2007; Alam et al., 2010; Medema et al., 2010, 2011a; Chindelevitch et al., 2012; Licona-Cassani et al., 2012). Even more surprising is the fact that pathway optimization of actinomycete metabolism has only been achieved at the level of precursor supply (Borodina et al., 2005b, 2008; Licona-Cassani et al., 2012). In such approaches, optimal solutions are found because there is congruence between cellular (maximize growth rate) and engineering objectives (maximize productivity). In order to properly optimize production of non-growth-associated metabolites (i.e., BNP) novel algorithms and new objective functions are to be incorporated to current protocols.

The last few years have seen the emergence of network reconstruction beyond metabolism. These next-generation network reconstructions account for expression coupled to metabolism

and even transcriptional regulation. The first such model, known as ME-model, was developed for *Thermotoga maritima* (Zhang et al., 2009; Lerman et al., 2012) rapidly followed by the *Escherichia coli* ME-model (O'Brien et al., 2013). Incorporation of gene expression in the mathematical framework allows these models to expand their predictive capabilities, which may be what is needed to model non-growth-associated metabolites such as BNPs. It is expected that as ME-models for actinomycetes become available, just like GEMs became available 10 years ago, multi-omics integration may be possible, and with it, models become more predictive.

## FUTURE DIRECTIONS

Like in model organisms, such as yeast and *E. coli*, systems biology in actinomycetes has immensely advanced our understanding of

this complex and fascinating bacterial family, offering insightful information regarding gene discovery, gene regulation, and pathway manipulation. The tools are highly developed and readily available yet integration and data analysis remain our main challenge. Just like finding a needle in a haystack, finding a key gene in a sea of data is extremely challenging; the main challenge remains the lack of tools for integration and visualization of large datasets. As such, systems biology in actinomycetes has yet to deliver real advances for the production of BNPs and discovery of novel bioactive compounds.

## FUNDING

We would like to thank The Queensland Government for the Accelerate fellowship to E.M. and The University of Queensland, the Australian Research Council, and the Mexican Council for Science and Technology CONACYT for financial support.

## REFERENCES

- Aigle, B., Lautru, S., Spiteller, D., Dickschat, J., Challis, G., Leblond, P., et al. (2014). Genome mining of *Streptomyces ambofaciens*. *J. Ind. Microbiol. Biotechnol.* 41, 251–263. doi:10.1007/s10295-013-1379-y
- Alam, M. T., Merlo, M., Consortium, T. S., Hodgson, D., Wellington, E., Takano, E., et al. (2010). Metabolic modeling and analysis of the metabolic switch in *Streptomyces coelicolor*. *BMC Genomics* 11:202. doi:10.1186/1471-2164-11-202
- Armengaud, J. (2010). Proteogenomics and systems biology: quest for the ultimate missing parts. *Expert Rev. Proteomics* 7, 65–77. doi:10.1586/Epr.09.104
- Aziz, R., Bartels, D., Best, A., Dejongh, M., Disz, T., Edwards, R., et al. (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi:10.1186/1471-2164-9-75
- Barona-Gomez, F., Cruz-Morales, P., and Noda-Garcia, L. (2012). What can genome-scale metabolic network reconstructions do for prokaryotic systematics? *Antonie Van Leeuwenhoek* 101, 35–43. doi:10.1007/s10482-011-9655-1
- Bentley, S., Chater, K., Cerdeno-Tarraga, A.-M., Challis, G., Thomson, N., James, K., et al. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417, 141–147. doi:10.1038/417141a
- Beste, D., Hooper, T., Stewart, G., Bonde, B., Avignone-Rossa, C., Bushell, M., et al. (2007). GSMN-TB: a web-based genome-scale network model of *Mycobacterium tuberculosis* metabolism. *Genome Biol.* 8, R89. doi:10.1186/gb-2007-8-5-r89
- Blin, K., Medema, M. H., Kazempour, D., Fischbach, M. A., Breitling, R., Takano, E., et al. (2013). antiSMASH 2.0: a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* 41, W204–W212. doi:10.1093/nar/gkt449
- Bobek, J., Strakova, E., Zikova, A., and Vohradsky, J. (2014). Changes in activity of metabolic and regulatory pathways during germination of *S. coelicolor*. *BMC Genomics* 15:1173. doi:10.1186/1471-2164-15-1173
- Borodina, I., Krabben, P., and Nielsen, J. (2005a). Genome-scale analysis of *Streptomyces coelicolor* A3(2) metabolism. *Genome Res.* 15, 820–829. doi:10.1101/gr.3364705
- Borodina, I., Schöller, C., Eliasson, A., and Nielsen, J. (2005b). Metabolic network analysis of *Streptomyces tenebrarius*, a *Streptomyces* species with an active Entner-Doudoroff pathway. *Appl. Environ. Microbiol.* 71, 2294–2302. doi:10.1128/aem.71.5.2294-2302.2005
- Borodina, I., Siebring, J., Zhang, J., Smith, C. P., Van Keulen, G., Dijkhuizen, L., et al. (2008). Antibiotic overproduction in *Streptomyces coelicolor* A3(2) mediated by phosphofructokinase deletion. *J. Biol. Chem.* 283, 25186–25199. doi:10.1074/jbc.M803105200
- Castellana, N., and Bafna, V. (2010). Proteogenomics to discover the full coding content of genomes: a computational perspective. *J. Proteomics* 73, 2124–2135. doi:10.1016/j.jprot.2010.06.007
- Chindelevitch, L., Stanley, S., Hung, D., Regev, A., and Berger, B. (2012). MetaMerge: scaling up genome-scale metabolic reconstructions with application to *Mycobacterium tuberculosis*. *Genome Biol.* 13, r6. doi:10.1186/gb-2012-13-1-r6
- Cimermancic, P., Medema, M. H., Claesen, J., Kurita, K., Wieland Brown, L. C., Mavrommatis, K., et al. (2014). Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* 158, 412–421. doi:10.1016/j.cell.2014.06.034
- Demain, A. (2014). Importance of microbial BNP and the need to revitalize their discovery. *J. Ind. Microbiol. Biotechnol.* 41, 185–201. doi:10.1007/s10295-013-1325-z
- Doroghazi, J., and Metcalf, W. (2013). Comparative genomics of actinomycetes with a focus on natural product biosynthetic genes. *BMC Genomics* 14:611. doi:10.1186/1471-2164-14-611
- Doroghazi, J. R., Albright, J. C., Goering, A. W., Ju, K.-S., Haines, R. R., Tchalukov, K. A., et al. (2014). A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat. Chem. Biol.* 10, 963–968. doi:10.1038/nchembio.1659
- Edwards, J. S., and Palsson, B. O. (1999). Systems properties of the *Haemophilus influenzae* metabolic genotype. *J. Biol. Chem.* 274, 17410–17416.
- Edwards, J. S., and Palsson, B. O. (2000). The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proc. Natl. Acad. Sci. U.S.A.* 97, 5528–5533. doi:10.1073/pnas.97.10.5528
- El-Enshasy, H. A., Mohamed, N. A., Farid, M. A., and El-Diwan, A. I. (2008). Improvement of erythromycin production by *Saccharopolyspora erythraea* in molasses based medium through cultivation medium optimization. *Bioresour. Technol.* 99, 4263–4268. doi:10.1016/j.biortech.2007.08.050
- Flardh, K., and Buttner, M. J. (2009). *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. *Nat. Rev. Microbiol.* 7, 36–49. doi:10.1038/nrmicro1968
- Gallien, S., Perrodou, E., Carapito, C., Deshayes, C., Reytrat, J. M., Van Dorsselaer, A., et al. (2009). Ortho-proteogenomics: multiple proteomes investigation through orthology and a new MS-based protocol. *Genome Res.* 19, 128–135. doi:10.1101/gr.081901.108
- Gupta, N., Tanner, S., Jaitly, N., Adkins, J. N., Lipton, M., Edwards, R., et al. (2007). Whole proteome analysis of post-translational modifications: applications of mass-spectrometry for proteogenomic annotation. *Genome Res.* 17, 1362–1377. doi:10.1101/Gr.6427907
- Hamed, J., Malekzadeh, F., and Saghati-Nia, A. E. (2004). Enhancing of erythromycin production by *Saccharopolyspora erythraea* with common and uncommon oils. *J. Ind. Microbiol. Biotechnol.* 31, 447–456. doi:10.1007/s10295-004-0166-1
- Henry, C. S., Dejongh, M., Best, A. A., Frybarger, P. M., Lindsay, B., and Stevens, R. L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* 28, 977–982. doi:10.1038/nbt.1672
- Hertweck, C. (2009). The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed.* 48, 4688–4716. doi:10.1002/anie.200806121

- Hesketh, A., Bucca, G., Laing, E., Flett, F., Hotchkiss, G., Smith, C., et al. (2007). New pleiotropic effects of eliminating a rare tRNA from *Streptomyces coelicolor*, revealed by combined proteomic and transcriptomic analysis of liquid cultures. *BMC Genomics* 8:261. doi:10.1186/1471-2164-8-261
- Huang, D., Li, Z.-H., You, D., Zhou, Y., and Ye, B.-C. (2015). Lysine acetylproteome analysis suggests its roles in primary and secondary metabolism in *Saccharopolyspora erythraea*. *Appl. Microbiol. Biotechnol.* 99, 1399–1413. doi:10.1007/s00253-014-6144-2
- Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., et al. (2003). Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat. Biotechnol.* 21, 526–531. doi:10.1038/nbt820
- Ikeda, H., Shin-Ya, K., and Omura, S. (2014). Genome mining of the *Streptomyces avermitilis* genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters. *J. Ind. Microbiol. Biotechnol.* 41, 233–250. doi:10.1007/s10295-013-1327-x
- Imam, S., Schäuble, S., Brooks, A. N., Baliga, N. S., and Price, N. D. (2015). Data-driven integration of genome-scale regulatory and metabolic network models. *Front. Microbiol.* 6:409. doi:10.3389/fmicb.2015.00409
- Jamshidi, N., and Palsson, B. (2007). Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the in silico strain iNJ661 and proposing alternative drug targets. *BMC Syst. Biol.* 1:26. doi:10.1186/1752-0509-1-26
- Jenke-Kodama, H., and Dittmann, E. (2009). Bioinformatic perspectives on NRPS/PKS megasynthases: advances and challenges. *Nat. Prod. Rep.* 26, 874–883. doi:10.1039/b810283j
- Jung, W., Yoo, Y., Park, J., Park, S., Han, A., Ban, Y., et al. (2011). A combined approach of classical mutagenesis and rational metabolic engineering improves rapamycin biosynthesis and provides insights into methylmalonyl-CoA precursor supply pathway in *Streptomyces hygroscopicus* ATCC 29253. *Appl. Microbiol. Biotechnol.* 91, 1389–1397. doi:10.1007/s00253-011-3348-6
- Khalidi, N., Seifuddin, F. T., Turner, G., Haft, D., Nierman, W. C., Wolfe, K. H., et al. (2010). SMURF: genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol.* 47, 736–741. doi:10.1016/j.fgb.2010.06.003
- Kim, D.-W., Chater, K., Lee, K.-J., and Hesketh, A. (2005). Changes in the extracellular proteome caused by the absence of the bldA gene product, a developmentally significant tRNA, reveal a new target for the pleiotropic regulator adpA in *Streptomyces coelicolor*. *J. Bacteriol.* 187, 2957–2966. doi:10.1128/jb.187.9.2957-2966.2005
- Kim, T. Y., Kim, H. U., and Lee, S. Y. (2010). Metabolite-centric approaches for the discovery of antibacterials using genome-scale metabolic networks. *Metab. Eng.* 12, 105–111. doi:10.1016/j.ymben.2009.05.004
- Koglin, A., and Walsh, C. T. (2009). Structural insights into nonribosomal peptide enzymatic assembly lines. *Nat. Prod. Rep.* 26, 987–1000. doi:10.1039/b904543k
- Lerman, J. A., Hyduke, D. R., Latif, H., Portnoy, V. A., Lewis, N. E., Orth, J. D., et al. (2012). In silico method for modelling metabolism and gene product expression at genome scale. *Nat. Commun.* 3, 929. doi:10.1038/ncomms1928
- Li, M. H., Ung, P. M., Zajkowski, J., Garneau-Tsodikova, S., and Sherman, D. H. (2009). Automated genome mining for natural products. *BMC Bioinformatics* 10:185. doi:10.1186/1471-2105-10-185
- Licona-Cassani, C., Lim, S., Marcellin, E., and Nielsen, L. K. (2014). Temporal dynamics of the *Saccharopolyspora erythraea* phosphoproteome. *Mol. Cell. Proteomics* 13, 1219–1230. doi:10.1074/mcp.M113.033951
- Licona-Cassani, C., Marcellin, E., Quek, L.-E., Jacob, S., and Nielsen, L. (2012). Reconstruction of the *Saccharopolyspora erythraea* genome-scale model and its use for enhancing erythromycin production. *Antonie Van Leeuwenhoek* 102, 493–502. doi:10.1007/s10482-012-9783-2
- Manteca, A., Fernandez, M., and Sanchez, J. (2005). Mycelium development in *Streptomyces antibioticus* ATCC11891 occurs in an orderly pattern which determines multiphase growth curves. *BMC Microbiol.* 5:51. doi:10.1186/1471-2180-5-51
- Manteca, A., Jung, H. R., Schwämmle, V., Jensen, O. N., and Sanchez, J. (2010). Quantitative proteome analysis of *Streptomyces coelicolor* nonsporulating liquid cultures demonstrates a complex differentiation process comparable to that occurring in sporulating solid cultures. *J. Proteome Res.* 9, 4801–4811. doi:10.1021/pr100513p
- Manteca, A., Ye, J., Sanchez, J., and Jensen, O. N. (2011). Phosphoproteome analysis of *Streptomyces* development reveals extensive protein phosphorylation accompanying bacterial differentiation. *J. Proteome Res.* 10, 5481–5492. doi:10.1021/pr200762y
- Marcellin, E., Licona-Cassani, C., Mercer, T., Palfreyman, R., and Nielsen, L. (2013a). Re-annotation of the *Saccharopolyspora erythraea* genome using a systems biology approach. *BMC Genomics* 14:699. doi:10.1186/1471-2164-14-699
- Marcellin, E., Mercer, T. R., Licona-Cassani, C., Palfreyman, R. W., Dinger, M. E., Steen, J. A., et al. (2013b). *Saccharopolyspora erythraea*'s genome is organised in high-order transcriptional regions mediated by targeted degradation at the metabolic switch. *BMC Genomics* 14:15. doi:10.1186/1471-2164-14-15
- Medema, M. H., Alam, M. T., Heijne, W. H. M., Van Den Berg, M. A., Müller, U., Trefzer, A., et al. (2011a). Genome-wide gene expression changes in an industrial clavulanic acid overproduction strain of *Streptomyces clavuligerus*. *Microb. Biotechnol.* 4, 300–305. doi:10.1111/j.1751-7915.2010.00226.x
- Medema, M. H., Blin, K., Cimermancic, P., De Jager, V., Zakrzewski, P., Fischbach, M. A., et al. (2011b). antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* 39, W339–W346. doi:10.1093/nar/gkr466
- Medema, M. H., and Fischbach, M. A. (2015). Computational approaches of natural product discovery. *Nat. Chem. Biol.* 11, 639–648. doi:10.1038/nchembio.1884
- Medema, M. H., Trefzer, A., Kovalchuk, A., Van Den Berg, M., Müller, U., Heijne, W., et al. (2010). The sequence of a 1.8-Mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. *Genome Biol. Evol.* 2, 212–224. doi:10.1093/gbe/evq013
- Nett, M., Ikeda, H., and Moore, B. S. (2009). Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Nat. Prod. Rep.* 26, 1362–1384. doi:10.1039/b817069j
- Nielsen, P., and Krogh, A. (2005). Large-scale prokaryotic gene prediction and comparison to genome annotation. *Bioinformatics* 21, 4322–4329. doi:10.1093/bioinformatics/bti701
- Nieselt, K., Battke, F., Herbig, A., Bruheim, P., Wentzel, A., Jakobsen, O., et al. (2010). The dynamic architecture of the metabolic switch in *Streptomyces coelicolor*. *BMC Genomics* 11:10. doi:10.1186/1471-2164-11-10
- Novotna, J., Vohradsky, J., Berndt, P., Gramajo, H., Langen, H., Li, X.-M., et al. (2003). Proteomic studies of diauxic lag in the differentiating prokaryote *Streptomyces coelicolor* reveal a regulatory network of stress-induced proteins and central metabolic enzymes. *Mol. Microbiol.* 48, 1289–1303. doi:10.1046/j.1365-2958.2003.03529.x
- Oberhardt, M. A., Palsson, B. O., and Papin, J. A. (2009). Applications of genome-scale metabolic reconstructions. *Mol. Syst. Biol.* 5, 320. doi:10.1038/msb.2009.77
- O'Brien, E. J., Lerman, J. A., Chang, R. L., Hyduke, D. R., and Palsson, B. O. (2013). Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol. Syst. Biol.* 9, 693. doi:10.1038/msb.2013.52
- O'Brien, E. J., Monk, J. M., and Palsson, B. O. (2015). Using genome-scale models to predict biological capabilities. *Cell* 161, 971–987. doi:10.1016/j.cell.2015.05.019
- Ohnishi, Y., Ishikawa, J., Hara, H., Suzuki, H., Ikenoya, M., Ikeda, H., et al. (2008). Genome sequence of the streptomycin-producing microorganism *Streptomyces griseus* IFO 13350. *J. Bacteriol.* 190, 4050–4060. doi:10.1128/jb.00204-08
- Olano, C., Lombó, F., Méndez, C., and Salas, J. A. (2008). Improving production of bioactive secondary metabolites in actinomycetes by metabolic engineering. *Metab. Eng.* 10, 281–292. doi:10.1016/j.ymben.2008.07.001
- Oliynyk, M., Samborsky, M., Lester, J. B., Mironenko, T., Scott, N., Dickens, S., et al. (2007). Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL23338. *Nat. Biotechnol.* 25, 447–453. doi:10.1038/nbt1297
- Park, J. H., Lee, K. H., Kim, T. Y., and Lee, S. Y. (2007). Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 7797–7802. doi:10.1073/pnas.0702609104
- Piette, A., Derouaux, A., Gerkens, P., Noens, E. E. E., Mazzucchi, G., Vion, S., et al. (2005). From dormant to germinating spores of *Streptomyces coelicolor* A3(2): New perspectives from the crp null mutant. *J. Proteome Res.* 4, 1699–1708. doi:10.1021/pr050155b
- Qiu, Y., Cho, B.-K., Park, Y. S., Lovley, D., Palsson, B. O., and Zengler, K. (2010). Structural and operational complexity of the *Geobacter sulfurreducens* genome. *Genome Res.* 20, 1304–1311. doi:10.1101/gr.107540.110
- Reeves, A., Brikun, I., Cernota, W., Leach, B., Gonzalez, M., and Weber, J. (2006). Effects of methylmalonyl-CoA mutase gene knockouts on erythromycin production in carbohydrate-based and oil-based fermentations of *Saccharopolyspora erythraea*. *J. Ind. Microbiol. Biotechnol.* 33, 600–609. doi:10.1007/s10295-006-0094-3



- Reeves, A. R., Brikun, I. A., Cernota, W. H., Leach, B. I., Gonzalez, M. C., and Mark Weber, J. (2007). Engineering of the methylmalonyl-CoA metabolite node of *Saccharopolyspora erythraea* for increased erythromycin production. *Metab. Eng.* 9, 293–303. doi:10.1016/j.ymben.2007.02.001
- Rodríguez-García, A., Barreiro, C., Santos-Beneit, F., Sola-Landa, A., and Martín, J. F. (2007). Genome-wide transcriptomic and proteomic analysis of the primary response to phosphate limitation in *Streptomyces coelicolor* M145 and in a  $\Delta$ phoP mutant. *Proteomics* 7, 2410–2429. doi:10.1002/pmic.200600883
- Schilling, C. H., Covert, M. W., Famili, I., Church, G. M., Edwards, J. S., and Palsson, B. O. (2002). Genome-scale metabolic model of *Helicobacter pylori* 26695. *J. Bacteriol.* 184, 4582–4593. doi:10.1128/jb.184.16.4582-4593.2002
- Starcevic, A., Zucko, J., Simunkovic, J., Long, P. F., Cullum, J., and Hranueli, D. (2008). ClustScan: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures. *Nucleic Acids Res.* 36, 6882–6892. doi:10.1093/nar/gkn685
- Tanaka, Y., Komatsu, M., Okamoto, S., Tokuyama, S., Kaji, A., Ikeda, H., et al. (2009). Antibiotic overproduction by *rpsL* and *rsmG* mutants of various actinomycetes. *Appl. Environ. Microbiol.* 75, 4919–4922. doi:10.1128/AEM.00681-09
- Venter, E., Smith, R. D., and Payne, S. H. (2011). Proteogenomic analysis of bacteria and archaea: a 46 organism case study. *PLoS ONE* 6:e27587. doi:10.1371/journal.pone.0027587
- Walsh, C. T., and Fischbach, M. A. (2010). BNP version 2.0: connecting genes to molecules. *J. Am. Chem. Soc.* 132, 2469–2493. doi:10.1021/ja909118a
- Walsh, C. T., Garneau-Tsodikova, S., and Howard-Jones, A. R. (2006). Biological formation of pyrroles: nature's logic and enzymatic machinery. *Nat. Prod. Rep.* 23, 517–531. doi:10.1039/b605245m
- Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Brucoleri, R., et al. (2015). antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 43, W237–W243. doi:10.1093/nar/gkv437
- Weber, T., Rausch, C., Lopez, P., Hoof, I., Gaykova, V., Huson, D. H., et al. (2009). CLUSEAN: a computer-based framework for the automated analysis of bacterial secondary metabolite biosynthetic gene clusters. *J. Biotechnol.* 140, 13–17. doi:10.1016/j.jbiotec.2009.01.007
- Yagüe, P., López-García, M. T., Rioseras, B., Sánchez, J., and Manteca, Á. (2013a). Pre-sporulation stages of *Streptomyces* differentiation: state-of-the-art and future perspectives. *FEMS Microbiol. Lett.* 342, 79–88. doi:10.1111/1574-6968.12128
- Yagüe, P., Rodríguez-García, A., López-García, M. T., Martín, J. F., Rioseras, B., Sánchez, J., et al. (2013b). Transcriptomic analysis of *Streptomyces coelicolor* differentiation in solid sporulating cultures: first compartmentalized and second multinucleated mycelia have different and distinctive transcriptomes. *PLoS ONE* 8:e60665. doi:10.1371/journal.pone.0060665
- Zerikly, M., and Challis, G. L. (2009). Strategies for the discovery of new BNP by genome mining. *Chembiochem* 10, 625–633. doi:10.1002/cbic.200800389
- Zhang, Y., Thiele, I., Weekes, D., Li, Z., Jaroszewski, L., Ginalska, K., et al. (2009). Three-dimensional structural view of the central metabolic network of *Thermotoga maritima*. *Science* 325, 1544–1549. doi:10.1126/science.1174671
- Zou, X., Hang, H.-F., Chu, J., Zhuang, Y.-P., and Zhang, S.-L. (2009). Oxygen uptake rate optimization with nitrogen regulation for erythromycin production and scale-up from 50 L to 372 m<sup>3</sup> scale. *Bioresour. Technol.* 100, 1406–1412. doi:10.1016/j.biortech.2008.09.017

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Licona-Cassani, Cruz-Morales, Manteca, Barona-Gomez, Nielsen and Marcellin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Systems Biology of Microbial Exopolysaccharides Production

Ozlem Ates\*

*Department of Medical Services and Techniques, Nisantasi University, Istanbul, Turkey*

## OPEN ACCESS

### Edited by:

Alvaro R. Lara,  
Universidad Autónoma  
Metropolitana-Cuajimalpa, Mexico

### Reviewed by:

Alessandro Giuliani,  
Istituto Superiore di Sanità, Italy  
Adelfo Escalante,  
Universidad Nacional Autónoma de  
México, Mexico

### \*Correspondence:

Ozlem Ates  
ozlem.ates@nisantasi.edu.tr

### Specialty section:

This article was submitted to  
Systems Biology,  
a section of the journal  
Frontiers in Bioengineering and  
Biotechnology

**Received:** 20 August 2015

**Accepted:** 30 November 2015

**Published:** 18 December 2015

### Citation:

Ates O (2015) Systems Biology of  
Microbial Exopolysaccharides  
Production.  
Front. Bioeng. Biotechnol. 3:200.  
doi: 10.3389/fbioe.2015.00200

Exopolysaccharides (EPSs) produced by diverse group of microbial systems are rapidly emerging as new and industrially important biomaterials. Due to their unique and complex chemical structures and many interesting physicochemical and rheological properties with novel functionality, the microbial EPSs find wide range of commercial applications in various fields of the economy such as food, feed, packaging, chemical, textile, cosmetics and pharmaceutical industry, agriculture, and medicine. EPSs are mainly associated with high-value applications, and they have received considerable research attention over recent decades with their biocompatibility, biodegradability, and both environmental and human compatibility. However, only a few microbial EPSs have achieved to be used commercially due to their high production costs. The emerging need to overcome economic hurdles and the increasing significance of microbial EPSs in industrial and medical biotechnology call for the elucidation of the interrelations between metabolic pathways and EPS biosynthesis mechanism in order to control and hence enhance its microbial productivity. Moreover, a better understanding of biosynthesis mechanism is a significant issue for improvement of product quality and properties and also for the design of novel strains. Therefore, a systems-based approach constitutes an important step toward understanding the interplay between metabolism and EPS biosynthesis and further enhances its metabolic performance for industrial application. In this review, primarily the microbial EPSs, their biosynthesis mechanism, and important factors for their production will be discussed. After this brief introduction, recent literature on the application of omics technologies and systems biology tools for the improvement of production yields will be critically evaluated. Special focus will be given to EPSs with high market value such as xanthan, levan, pullulan, and dextran.

**Keywords:** EPS, microbial production, exopolysaccharides, systems biology, xanthan, levan, pullulan, dextran

## INTRODUCTION

Biopolymer (also called renewable polymers) is used as a term to describe polymers produced by biological systems and polymers that are not synthesized chemically but are derived from biological starting materials such as amino acids, sugars, and natural fats (Tang et al., 2012). Consequently, biopolymers can be classified as synthetic or natural polymers (Vroman and Tighzert, 2009). The biopolymers are superior to petrochemical-derived polymers in several aspects that include biocompatibility, biodegradability, and both environmental and human compatibility. Although petroleum-based polymers have negative effects to environment and humanity such as toxicity, defiance to biodegradation, and waste accumulation, they have been used in a variety of industrial

applications. Therefore, in response to these problems, biopolymers are a suitable alternative that the researchers were looking (Keshavarz and Roy, 2010).

Today, several microorganisms are identified as microbial biopolymer producers and these polymers can be found as attached to the cell surface or extracted from the fermentation medium. Bacteria use these microbial biopolymers as storage materials in response to particular environmental stresses (Sanchez-Garcia et al., 2010). Due to their biological functions microbial polysaccharides can be generally classified as intracellular storage polysaccharides (glycogen), capsular polysaccharides (e.g., K30 O-Antigen), and extracellular bacterial polysaccharides (for example, levan, xanthan, sphingane, alginate, pullulan, cellulose, etc.), which are important for biofilm formation and pathogenicity (Schmid and Sieber, 2015).

Microbial polysaccharides that are produced by microorganisms and secreted out of the cell are defined as exopolysaccharides (EPSs). In nature, they have a significant role for protection of the cell, adhesion of bacteria to solid surfaces, and participating in cell-to-cell interactions (Nicolaus et al., 2010). In recent years, there is a significant interest on microbial EPSs since they have different structural and functional properties (Morris and Harding, 2009). EPSs are important resources for hydrocolloids used in food, pharmaceutical, chemical, and many other industries (Ahmad et al., 2015). Due to their many interesting physicochemical and rheological properties with novel functionality, the microbial EPSs act as new biomaterials and find a wide range of applications in many industrial sectors such as textiles, detergents, adhesives, microbial enhanced oil recovery (MEOR), wastewater treatment, dredging, brewing, downstream processing, cosmetology, pharmacology, and food additives (Rühmann et al., 2015). Xanthan, dextran, and pullulan are examples of microbial polysaccharides with a considerable market due to their exceptional properties. However, plant and algal polysaccharides such as starch, galactomannans, pectin, carrageenan, and alginate still include a major part of the hydrocolloid market, which has a market value of 4 million US dollars in 2008, 3.9 billion US dollars in 2012, and this value is expected to reach 7 billion US dollars by 2019 (Williams et al., 2007; Patel and Prajapati, 2013). Since microbial EPSs enable fast and high yielding production processes under controlled conditions, they are economically competitive to the plant and algal origin polysaccharides, which are affected by climatological and geological environmental conditions (Kaur et al., 2014). Although there is an increased attraction for microbial EPSs in industrial and medical applications and they are related with high-value applications, only a few bacterial EPSs have achieved to be used commercially such as xanthan, gellan, and dextran due to high production costs (Freitas et al., 2011; Llamas et al., 2012).

Due to the exceptionally high production costs, microbial EPSs could never find their proper place in the polymer market and therefore, high-level EPSs producing microbial systems gain escalating industrial importance. Increasing significance of EPSs in industrial and medical biotechnology calls for the elucidation of the interrelations between metabolic pathways and biosynthesis mechanism in order to control and hence improve microbial productivity. Therefore, extensive interest has been dedicated to understand bacterial EPSs biosynthesis mechanism and pathways

and enhance productivity within the past years. Omics technologies such as genome sequencing, functional genomics, protein structure analysis, and new bioinformatics tools have been used to identify new EPS biosynthesis pathways and understand the principles of EPS formation (Schmid and Sieber, 2015).

A modeling approach that linked omics data and the simulation of variable expression and enzyme activity will provide information about a cell's macromolecular machinery (Lerman et al., 2012). For this purpose, genome-based and genome-scale metabolic reconstructions can be used to understand and predict phenotypes of a microbial species (Hanemaaijer et al., 2015). Therefore, a systems biology approach constitutes an important step toward understanding the interplay between metabolism and microbial EPS biosynthesis and further enhances its metabolic performance for industrial application. **Figure 1** demonstrated a brief summary of integration of omics studies with systems biology.

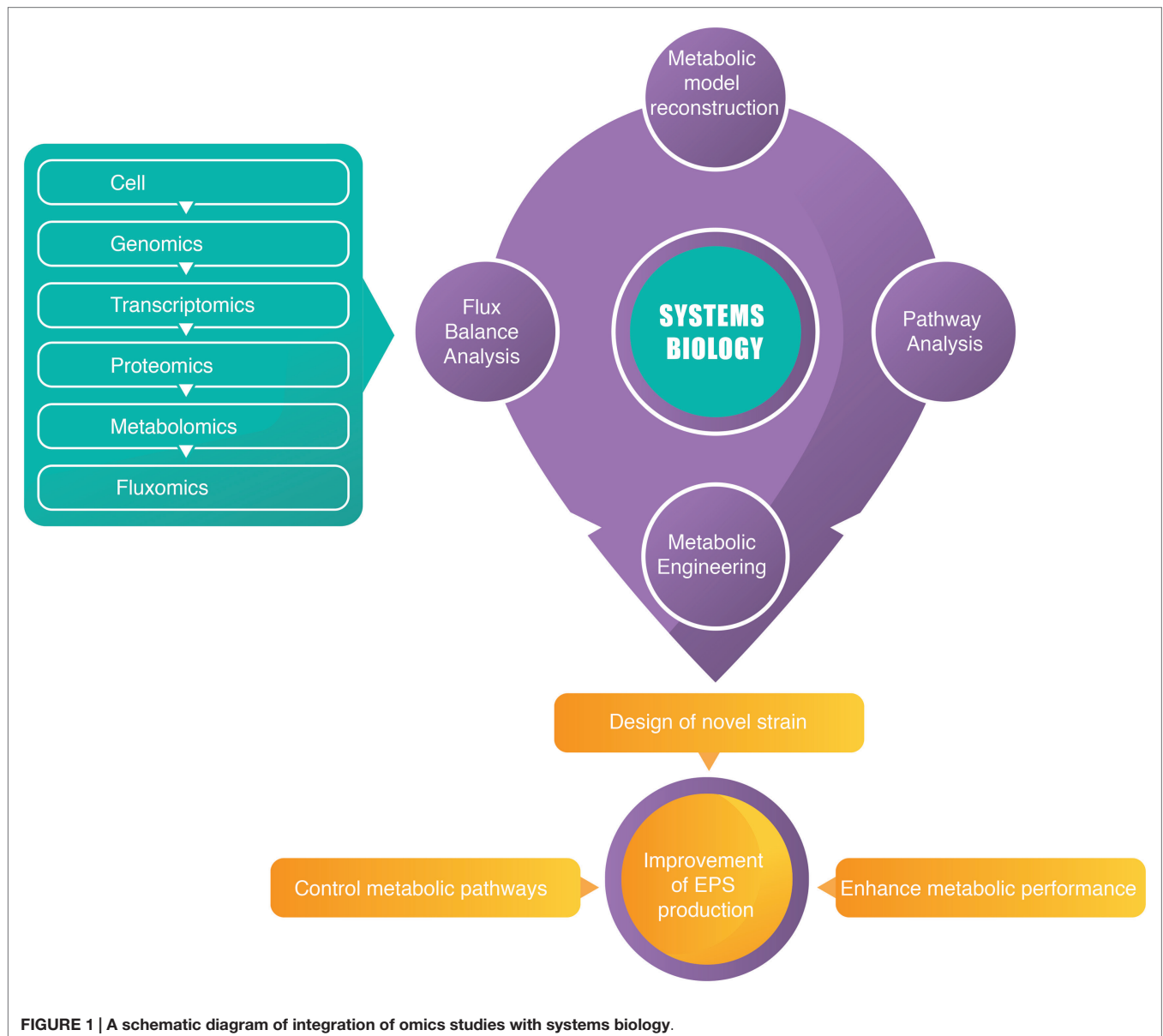
In this review after a brief description of the microbial EPSs, biosynthesis mechanism and important factors for their production, recent literature on the application of omics technologies and systems biology tools for the improvement of production yields will be critically evaluated. Microbial EPSs with high market value such as xanthan, dextran, scleroglucan, pullulan, and levan are specially focused.

## MICROBIAL EXOPOLYSACCHARIDES

The biopolymers produced by microorganisms were categorized to four main groups: polyesters, polyamides, inorganic polyanhydrides, and polysaccharides (Crescenzi and Dentini, 1996; Rehm, 2009, 2010). Since the microbial biopolymers serve as reserve material or as part of a protective mechanism, the biopolymer producer microorganisms have significant advantages under certain environmental conditions (Rehm, 2010). The first bacterial polymer dextran was discovered by Pasteur (1861) in the mid-nineteenth century as a microbial product in wine and the bacterium *Leuconostoc mesenteroides* was identified by Van Tieghem (1878) as dextran producer strain.

The bacterial polysaccharides that are synthesized and secreted by various microorganisms into the extracellular environment either as soluble or insoluble polymers are defined as EPSs. Their compositions, functions, chemical, and physical properties that establish their primary conformation vary from one bacterial species to another. EPSs are composed of mainly of carbohydrates (a wide range of sugar residues) and some non-carbohydrate substituents (such as acetate, pyruvate, succinate, and phosphate) (Vu et al., 2009; Nicolaus et al., 2010; Llamas et al., 2012; Staudt et al., 2012).

Most of EPSs producer bacteria have been described to produce either homo or heteropolysaccharide (Kumar et al., 2007). On the other hand, bacteria (*Serratia marcescens*, *Aeromonas salmonicida*, and *Pseudomonas* sp. strain NCIB 2021) that were able to produce two different polysaccharides have been reported (Kwon et al., 1994). Due to linkage bonds and nature of monomeric units, homopolysaccharides can be categorized as  $\alpha$ -D-glucans,  $\beta$ -D-glucans, fructans, and polygalactan. D-glucose, D-galactose, L-rhamnose, and N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), or glucuronic acid (GlcA)



are the repeating units of heteropolysaccharides and occasionally non-carbohydrate substituent such as phosphate, acetyl, and glycerol. Homopolysaccharides and heteropolysaccharides are also dissimilar in synthetic enzymes and site of synthesis. Biosynthesis of homopolysaccharides requires specific substrates like sucrose, while the residues of heteropolysaccharide are produced intracellularly and precursors are located across the membrane by isoprenoid glycosyl carrier lipids for extracellular polymerization (Nwodo et al., 2012). EPSs have also been classified in seven categories based on their functionality by Flemming and Wingender (2010) as constructive or structural (serve in the matrix help water retention and cell protection), sorptive (composed of charged polymers), surface-active (including molecules with amphiphilic behavior), active, informative, redox-active, and nutritive.

EPS affords self-protection for cells from desiccation, predation, the effects of antibiotics, antimicrobial substances, antibodies, bacteriophages and adherence to other bacteria, animal, and plant tissues under different stress conditions such as biotic stress, competition, and abiotic stresses, including temperature, light intensity, or pH (Mata et al., 2006; Kumar et al., 2007; Kumar and Mody, 2009; Ordax et al., 2010; Donot et al., 2012; Staudt et al., 2012). Additionally, EPS supplies bacterial aggregation, surface attachment, and symbiosis of plant-microbe; hence, it is a crucial property for wastewater treatment and soil aggregation. Furthermore, pathogenicity of a microorganism is related with the production of capsular EPS and depends on the rate and amount of EPS synthesis (Kumar et al., 2007).

Microorganisms are often linked with a high cellular density biofilm and its stability is controlled by EPSs through interactions

between the polysaccharide chains. Moreover, microbial diversity is biologically supported by EPS to constitute a substrate for the microbial growth. Biofilm formation plays a crucial role both in adhesion and in adaptation of bacteria to the physicochemical conditions of the environment (Donot et al., 2012).

Environmental factors and specific culture conditions such as pH, temperature, carbon-to-nitrogen (C/N) ratio, oxygenation rate, and carbon sources can impact EPS production. EPS composition can be altered by conditional changes (differing monosaccharides or by monosaccharide molar ratio). For instance, due to the carbon source, *Lactobacillus casei* has been shown to alter the chemical composition of its EPS (Staudt et al., 2012). Furthermore, production of microbial EPSs in bioreactors enables to optimize growth and production yields by studying physiology and genetic engineering (Delbarre-Ladrat et al., 2014).

Due to their unique and complex chemical structures that offers beneficial bioactive functions, biocompatibility, and biodegradability, microbial EPSs have found a wide range of application areas in chemical, food, pharmaceutical, cosmetics, packaging industries, agriculture, and medicine in which they can be used as adhesives, absorbents, lubricants, soil conditioners, cosmetic, drug delivery vehicles, textiles, high-strength materials, emulsifiers, viscosifiers, suspending, and chelating agents. In recent years, several novel bacterial EPSs have been isolated and identified; however, a few of them have achieved to have significant commercial value due to the high production costs (Mata et al., 2006; Kumar et al., 2007; Nicolaus et al., 2010; Freitas et al., 2011; Llamas et al., 2012; Delbarre-Ladrat et al., 2014). Bacterial EPSs such as xanthan, gellan, dextran, and curdlan with superior physical and chemical properties are used instead of plant (guar gum or pectin) or algae (e.g., carrageenan or alginate) polysaccharides in traditional applications (Kumar and Mody, 2009; Nicolaus et al., 2010; Freitas et al., 2011; Liang and Wang, 2015). Other bacterial EPSs such as levan, pullulan, and wellan with unique properties and biological activities have found new commercial opportunities (Freitas et al., 2011).

GalactoPol, which is synthesized by *Pseudomonas oleovorans* and composed mainly of galactose, and a fucose containing EPS FucoPol that is synthesized by *Enterobacter A47* have been recently reported bacterial EPSs with great commercial potential. Almost 30 species of lactic acid bacteria (LAB) are also known as polysaccharide producers and one of the commercial EPS dextran producer *Leuconostoc mesenteroides* is a LAB; however, low production yields avoid LAB species to be exploited commercially. Besides, *Lactobacilli* are GRAS (generally recognized as safe) bacteria and their EPS could be utilized in foods (Badel et al., 2011).

After the discovery of the various EPSs, the activities of enzymes related with EPS production were investigated and radioisotope-labeled precursors were used to elucidate the metabolic pathways for microbial biosynthesis. Moreover, understanding the molecular and regulatory mechanisms behind the biosynthesis of microbial polymers is an essential requirement for engineering bacteria leading to production of tailor-made biopolymers with high-value applications for industrial and medical applications with an economic cost (Rehm, 2009, 2010).

## BACTERIAL SYNTHESIS OF EXOPOLYSACCHARIDES

Extensive progress has been made in elucidating the biosynthetic and genetic mechanisms of biosynthesis of bacterial polysaccharides in recent years. The mechanism of biosynthesis and the precursors required illustrate diversity for different classes of EPSs. EPSs are synthesized by bacteria extracellularly or intracellularly (Boels et al., 2001; Kumar et al., 2007; Badel et al., 2011; Freitas et al., 2011; Li and Wang, 2012). Genes required for EPS production are responsible for encoding regulation, chain-length determination, repeat-unit assembly, polymerization, and export. The mechanism regulating EPS biosynthesis is a challenged topic to be understood despite accumulating knowledge of EPS gene organization (Péant et al., 2005). Regulation of EPS biosynthesis is related with various physiological and metabolic parameters such as the availability of sugar precursors and the expression level of enzymes (Delbarre-Ladrat et al., 2014). Information on genetics of certain EPS like xanthan is abundant; however, genetic data for other EPS synthesis (i.e., pullulan) is still limited (Donot et al., 2012).

Bacterial EPSs are mostly generated intracellularly and exported to the extracellular environment with the exception of homopolysaccharides such as dextran, levan, and mutan that are synthesized outside the cells by the action of secreted enzymes that convert the substrate into the polymer. The enzymes involved in EPS synthesis are found at different regions of the cell and can be characterized into four categories. The first group is intracellular enzymes such as hexokinase, which phosphorylates glucose (Glc) to glucose-6-phosphate (Glc-6-P). They are also involved in other cellular metabolisms. The second group is required to catalyze conversion of sugar nucleotides. Uridine-5'-diphosphate (UDP)-glucose pyrophosphorylase that catalyzes the conversion of Glc-1-P to UDP-Glc, which is one of the key molecules in EPS synthesis can be given as an example for this class of enzymes. Another enzyme group is glycosyltransferases (GTFs) that are located in the cell periplasmic membrane. The sugar nucleotides are transferred by GTFs to a repeating unit attached to glycosyl carrier lipid. The enzymatic functions, the structures, and identification of the genes that encode GTFs has been investigated intense and due to amino acid sequence similarities more than 94 GTF families were reported in the Carbohydrate-Active EnZymes (CAZy) database (<http://www.cazy.org>) (Li and Wang, 2012). The last class is presumably involved in the polymerization of the macromolecules and situated outside the cell membrane and the cell wall (Kumar et al., 2007).

The general mechanisms for the production of bacterial polysaccharides are Wzx/Wzy-dependent pathway, the ATP-binding cassette (ABC) transporter-dependent pathway, the synthase-dependent pathway, and the extracellular synthesis by use of a single sucrose protein. Inside the cell, the precursor molecules are transformed by enzymes and produce activated sugars/sugar acids in the first three mechanisms. Alternatively, in extracellular production pathway by direct addition of monosaccharides obtained by cleavage of di- or trisaccharides, the polymer strand is elongated (Schmid and Sieber, 2015). These

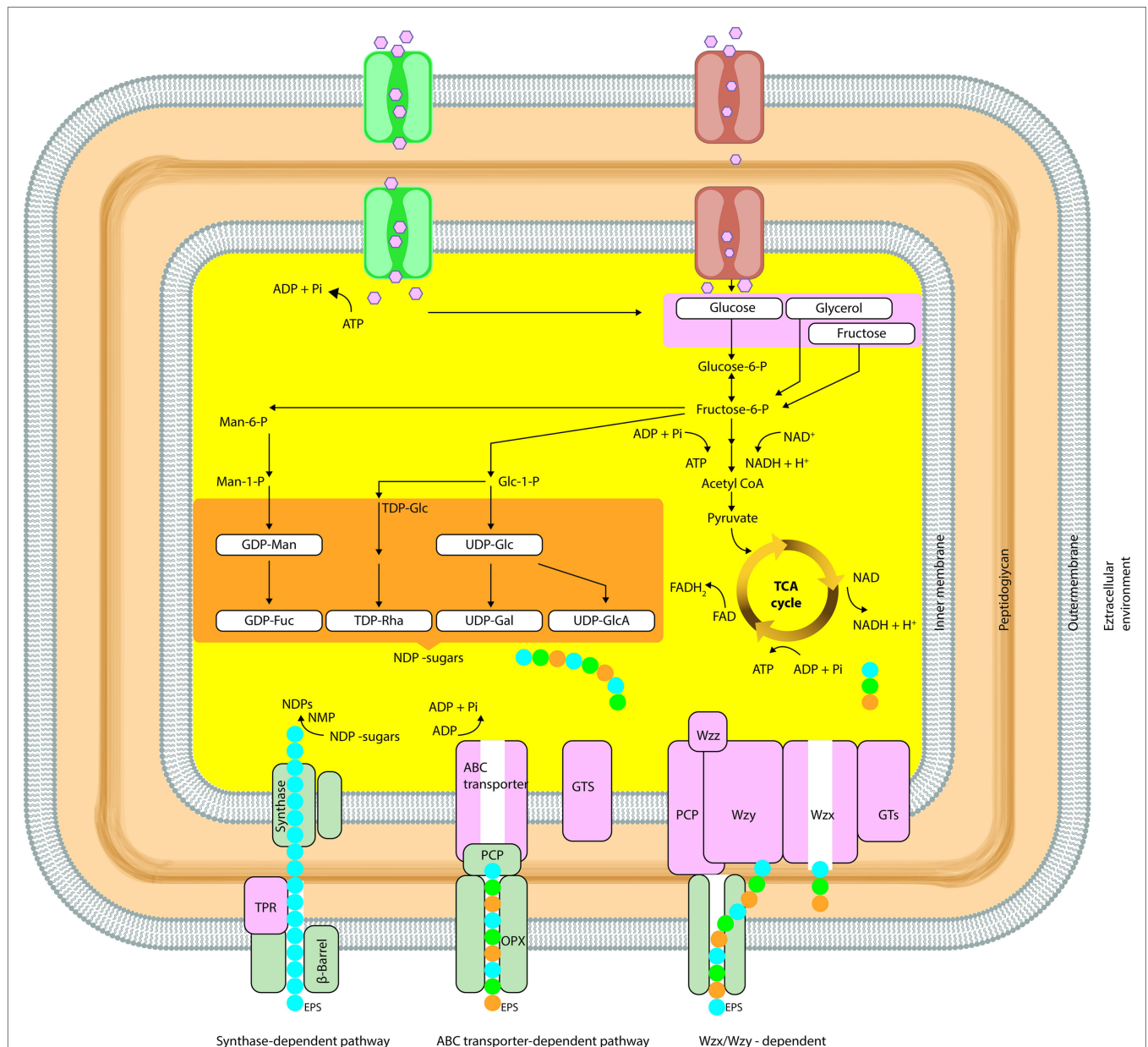


general mechanisms of EPS biosynthesis were demonstrated in detail in **Figure 2**.

In the Wzx/Wzy-dependent process, activated sugars are linked in a specific sequence to a lipid carrier by GTFs until the repeating unit is formed involving a Wzy protein. In the Wzx/Wzy-independent (ABC transporter-dependent) pathway, polymerization occurs at the cytoplasmic side of the inner membrane. The genes, which are required for high-level polymerization and surface assembly, are described as *wza* (encoding an outer-membrane protein), *wzb* (encoding an acid phosphatase), and *wzc* (encoding an inner-membrane tyrosine autokinase). In

most Gram-negative bacteria (i.e., *Erwinia* spp., *Methylobacillus* sp. strain 12S, *Rhizobium* spp., and *Xanthomonas campestris*), EPS biosynthesis and export have been reported to occur via the Wzx/Wzy-independent and Wzx/Wzy-dependent pathway (Arco et al., 2005; Cescutti et al., 2010; Freitas et al., 2011).

EPS secretion can occur in the presence or absence of a lipid acceptor molecule in the synthase-dependent pathway, which secretes complete polymer strands across the membranes and the cell wall, and is not dependent of a flippase for translocating repeat units. In this system, the polymerization and the translocation process are performed simultaneously by a single synthase



**FIGURE 2 | Simplified schematic diagram summarizing the biosynthetic pathways (Synthase dependent, ABC transporter, Wzx/Wzy dependent) involved in microbial EPS synthesis (OPX, outer-membrane polysaccharide; PCP, polysaccharide copolymerase; TPR, tetratricopeptide repeat proteins).**

protein, a membrane-embedded glycosyl transferase. These pathways are often utilized for the assembly of homopolymers requiring only one type of sugar precursor such as curdlan [ $\beta$ -(1-3)-linked glucose monomers] or bacterial cellulose [ $\beta$ -(1-4)-linked glucose units] (Rehm, 2010; Whitney and Howell, 2013; Schmid and Sieber, 2015). Regulation of polymerization is implemented by an inner-membrane receptor, in Gram-negative synthase-dependent secretion systems such as *P. aeruginosa* alginate and *Gluconacetobacter xylinus* cellulose (Whitney and Howell, 2013).

In extracellular synthesis, polymerization reaction occurs as transfer of a monosaccharide from a disaccharide to a growing polysaccharide chain in the extracellular environment. This type of production of EPSs is uncomplicated; independent of the central carbon metabolism besides there is a limited variation in structure. The extracellular EPS synthesis can occur for homopolysaccharides (dextran, levan, and mutan) by extracellular GTF (Boels et al., 2001; Finore et al., 2014).

The intracellular biosynthesis of homo- and heteropolysaccharides includes production of (ir)regular repeating units from sugar nucleotide precursors, which are also involved in the biosynthesis of several cell wall components and can therefore be considered essential for growth (Boels et al., 2001; Nicolaus et al., 2010; Li and Wang, 2012). Direct precursors for bacterial EPS biosynthesis are formed intracellularly from intermediates of the central carbon metabolism. The precursors and donor monomers for the biosynthesis of most repeating units are sugar nucleotides such as nucleoside diphosphate sugars (such as ADP-glucose), nucleoside diphosphate sugar acids (such as GDP-mannuronic acid), and nucleoside diphosphate sugar derivatives [such as UDP-glucose, UDP-N-acetyl glucosamine, UDP-galactose, and deoxythymidine diphosphate (dTDP)-rhamnose] (Barreto et al., 2005; Péant et al., 2005; Rehm, 2010). These sugars can be transported by basically three different ways: ATP hydrolysis coupled to sugar translocation via a sugar transport ATPase, the import with coupled to transport of ions and other solutes, and transport via the phosphoenolpyruvate (PEP) transport system (PTS) (Barreto et al., 2005; Péant et al., 2005).

GTFs (EC 2.4.x.y) catalyze heteropolysaccharide biosynthesis which has numerous intracellular steps and only the last step that polymerization of repeating units occurs is extracellular. Depending on substrate type, uptake of sugars is achieved through a passive or an active transport system by the cell in the first step. Subsequently, the substrate is catabolized in the cytoplasm through glycolysis and sugar nucleotides are formed. The biosynthesis of activated precursors [energy-rich monosaccharides, mainly nucleoside diphosphate sugars (NDP-sugars)], which are derived from phosphorylated sugars is occurred. Finally, EPS is secreted to extracellular environment therefore their secretion from cytoplasm through cell membrane without compromising the critical barrier properties is a challenging process (Badel et al., 2011; Freitas et al., 2011).

Conversely, homopolysaccharides are synthesized extracellularly by GTFs. This class is defined as glycansucrases class (E.C. 2.4.x.y) and dissimilar to classical Leloir-type GTF they utilize sucrose as donor substrate instead of nucleotide-sugars. The transfer of monosaccharides, generating a glycosidic bond,

from activated molecules to an acceptor molecule is catalyzed by these enzymes. Energy released by degradation of sugars is used to catalyze transfer of a glycosyl residue on forming polysaccharide. Due to the product of biosynthesis, the enzymes can be differentiated between transglucosylases (EC 2.4.1.y) and transfructosylases (EC 2.4.1.y or 2.y). Transglucosylases class includes dextransucrase, mutansucrase, and reuteransucrase (EC 2.4.1.5), which are high molecular weight extracellular enzymes and catalyze hydrolysis of sucrose to glucose and fructose and glucosyl transfer on carbohydrate or non-carbohydrate compounds. EPS structures can be varied based on different enzymes intervention and the synthesis of each polysaccharide is catalyzed by a specific GTF; therefore, two products encoded by two genes of *gtf* result in two different EPS. In addition, the enzyme conformation affects branching degree of homopolysaccharides. Levansucrases (EC 2.4.1.10) and inulosucrases (EC 2.4.1.9) from transfructosidases class produce levan and inulin type fructans. *Ftf* genes are induced under stress conditions and sucrose hydrolysis and fructosyl transfer on fructan polymerized chain or syntheses of tri- or tetrasaccharides are catalyzed. In fructan, glucose is the non-terminal reducing residue (G-Fn) (Badel et al., 2011).

## OMICS STUDIES AND SYSTEMS BIOLOGY AND OF MICROBIAL BIOPOLYMER AND MICROBIAL EPS PRODUCTION

Systems biology offers valuable application areas in molecular sciences, medicine, pharmacy, and engineering such as pathway-based biomarkers and diagnosis, systematic measurement and modeling of genetic interactions, systems biology of stem cells, identification of disease genes, drug design, strain development, bioprocess optimization (Medina-Cleghorn and Nomura, 2013).

Strain improvement using systems level analysis of metabolic, gene regulatory and signaling networks, and integration of omics data are the most focused subjects of systems biology. Biochemical and bioprocess engineering principles are applied for optimization of upstream-to-downstream bioprocesses at first stages of strain development (Barrett et al., 2006). Process development has been a supporter of the scientific achievements in systems biology, mostly in the areas of transcriptomics, proteomics, metabolomics, and fluxomics with availability of genome sequences for production organisms. The applications of systems biology in industry become a challenged subject (Otero and Nielsen, 2010).

In recent years, the enormous amounts of genome sequencing projects have resulted in accumulation of complete genome sequence information for a number of species. This information is valuable for understanding biological capabilities of organisms and biological processes such as signal transduction and cellular metabolism at the system levels. Therefore, an ever-increasing number of models for bacteria and more papers that describe new reconstruction tools and improvements have been published. Genome-scale constraint-based metabolic models have been reconstructed for several organisms and such constraint-based models can be quickly generated by software packages using an organism's genomic, biochemical, and physiological data. These

metabolic models have been used to integrate high-throughput data understand cellular metabolism, to develop metabolic engineering strategies, to design media and processes, to consider theoretical capabilities, and to control the process online, which illustrates its usefulness for development and optimization of process. Metabolic models are used for generating new hypotheses and targeting promising areas in engineering field. Metabolic engineering studies have been performed to modify microbes to produce industrially relevant biochemicals and biofuels such as ethanol and higher alcohols, fatty acids, amino acids, shikimate precursors, terpenoids, polyketides, and polymer precursors (e.g., 1,4-butanediol) (Henry et al., 2010; Baart and Martens, 2012; Xu et al., 2013; Long et al., 2015; Simeonidis and Price, 2015).

Genome-scale metabolic networks have great achievement in development of metabolic engineering strategies for strain improvement mainly in five industrial fields: food and nutrients, biopharmaceuticals, biopolymer materials, microbial biofuels, and bioremediation. Metabolic models are built to improve the yield of fermentation by products and explore metabolic mechanisms and processes in food and nutrients industry. Several biopharmaceuticals and the productivity of useful biopolymers and their precursors have improved by genome-scale metabolic model-guided metabolic engineering strategy (Xu et al., 2013). Dupont's near-decade long optimization of *Escherichia coli* for bioproduction of 1,3-propanediol is an important genome-scale metabolic engineering application (Nakamura and Whited, 2003). The industrially optimized strain required up to 26 genomic changes including insertions, deletions, and regulatory modifications. Recent advances in constraint-based modeling have enabled *in silico* prediction of genomic targets for the enhancement of strain performance or product yield (Esvelt and Wang, 2013). The engineering strategies have been successfully implemented for the improvements in the yield or production process, alterations in the degree of polymerization, removal of side chains or non-sugar substituents, or heterologous expression of EPS biosynthesis gene clusters (Becker, 2015). Additionally, the gene clusters of significant EPSs were figured out in **Figure 3**.

Due to their superior properties, wide application areas, there is a high demand to improve microbial EPS production with an economical cost. Therefore, the omics data and tools were utilized

to perform systems biology approaches to understand and control EPS biosynthesis mechanism, design novel strains, and enhance productivity.

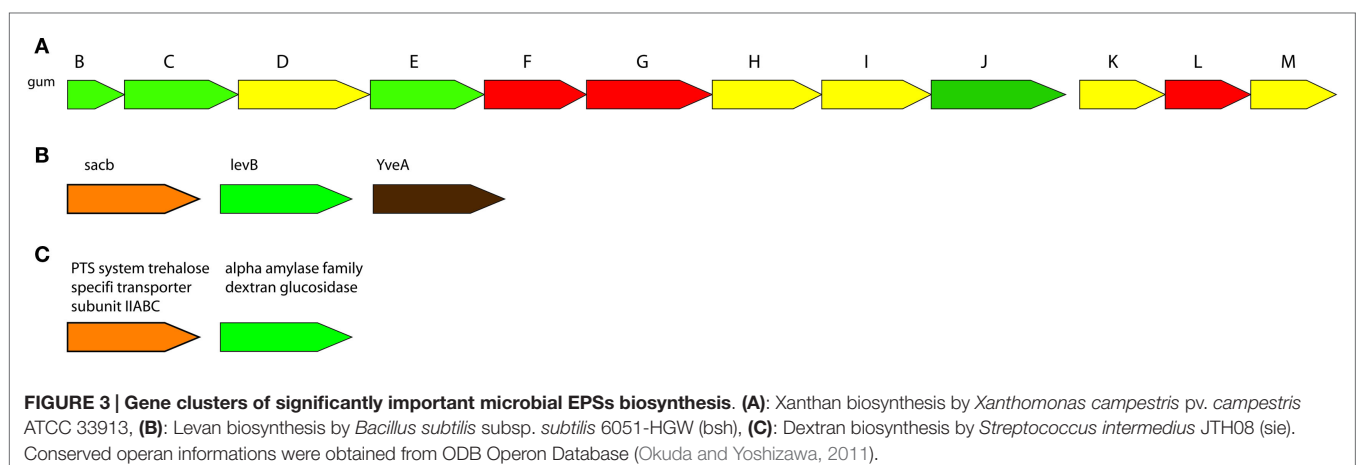
Natural or engineered microorganisms could synthesize many biopolymers and their monomers such as poly-3-hydroxyalkanoates (PHAs), polylactic acid (PLA), polysaccharides, carboxylic acids, and butanediols. Systems biology approach and genome-scale metabolic models-guided metabolic engineering strategies have been successfully employed to enhance the productivity of useful biopolymers and their precursors (Xu et al., 2013).

Jung et al. reported direct synthesis of PLA, which is a promising biomass-derived homopolymer and its copolymer, poly (3-hydroxybutyrate-co-lactate), P (3HB-co-LA), by direct fermentation of metabolically engineered *E. coli*. In typical conditions, PLA production involves two steps fermentative production of lactic acid followed by chemical polymerization with low production yields. In this study, *in silico* genome-scale metabolic flux analysis was performed to determine metabolic engineering targets to improve *E. coli* strain. The engineering process was achieved by knocking out the *ackA*, *ppc*, and *adhE* genes and by replacing the promoters of the *ldhA* and *acs* genes with the *trc* promoter, and therefore, an 11 wt% enhancement of PLA production was obtained.

Polyhydroxyalkanoates (PHA) synthesizing capacity of *Pseudomonas putida* was investigated by genome-scale metabolic model of this microorganism and survival under anaerobic stress was achieved by introducing the *ackA* gene from *Pseudomonas aeruginosa* and *Escherichia coli* (Sohn et al., 2010).

Cai et al. (2011) reported the draft genome of the moderately halophilic bacterium *Halomonas* sp. TD01. In this study, several genes relevant to PHA and osmolytes biosynthesis were analyzed providing invaluable clues for understanding of the evolution and genes transfer, the strategic guidance of the genetic engineering of halophilic *Halomonas* sp. TD01 for co-production of PHA and ectoine.

The analysis of the genes required for the synthesis of the EPS mauran by *H. maura* strain S-30 was performed to identify gene cluster in this strain. Three conserved genes, *epsA*, *epsB*, and *epsC*, also a *wzx* homolog, *epsJ*, which indicates that mauran is formed by a Wzy-dependent system, were found. It was also reported





that *eps* gene cluster reaches maximum activity during stationary phase, in the presence of high salt concentrations (5% w/v), which was investigated by transcriptional expression assays using a derivative of *H. maura* S-30, which carries an *epsA::lacZ* transcriptional fusion (Arco et al., 2005).

Generation of monomers including propanediols, butanediols, diamines, and terpenoids by microorganism through easier biosynthetic pathways was also reported (Lee et al., 2011, 2012; Curran and Alper, 2012). Additionally, production of monomers has been improved using systems biology approaches (Yim et al., 2011; Ng et al., 2012).

Genome-scale metabolic model-guided metabolic engineering approach has been performed and successfully used to improve production yields of various biopolymers and their precursors in synthetic material industry. The microbial production of monomers such as butanediols that are important raw materials in this industry are enhanced by genome-scale metabolic model strategies (Xu et al., 2013). For instance, Ng et al. (2012) have designed and constructed *S. cerevisiae* strains with improved production of 2,3-butanediol with gene deletion strategy, in which disruption of alcohol dehydrogenase (ADH) pathway is required, by performing *in silico* genome-scale metabolic analysis. Yim et al. (2011) have used biopathway prediction algorithm to elucidate possible pathways for 1,4-butanediol (BDO) biosynthesis. Strain development was performed by engineering the *E. coli* host to enhance anaerobic operation of the oxidative tricarboxylic acid cycle and drive the BDO pathway. The engineered strain was able to produce BDO from glucose, xylose, sucrose and biomass-derived mixed sugar streams. Furthermore, the productions of some important carboxylic acid monomers, used as raw materials in synthetic material industry, such as formic acid, malic acid, and succinic acid were improved in engineered *S. cerevisiae* or *E. coli* via genome-scale metabolic models-guided metabolic engineering strategies (Lee et al., 2005b; Wang et al., 2006; Moon et al., 2008; Kennedy et al., 2009). These successfully implemented studies will be helpful to improve microbial EPS production and to design industrial strategies.

Systems metabolic engineering of *E. coli* or *Corynebacterium glutamicum* as efficient cell factories has resulted in overproduction of 1,5-diaminopentane as building block for novel biopolymers (Kind and Wittmann, 2011). The importance of the Entner–Doudoroff pathway in PHB production was predicted by stoichiometric flux analysis of recombinant *E. coli* metabolic model and confirmed experimentally (Hong et al., 2003). The dynamics of PHA copolymer structure and properties were identified by mathematical models during its *in vivo* accumulation (Aldor and Keasling, 2003). Besides, an optimal carbon source switching strategy for the production of block copolymers was described by a population balance model in *Ralstonia eutropha* system (Mantzaris et al., 2001).

Previously genome-scale metabolic model reconstructions of biopolymer producer strain such as *Pseudomonas putida* (Nogales et al., 2008; Puchalka et al., 2008) and *Pseudomonas aeruginosa* (Oberhardt et al., 2008) have been published, as they could be used to elucidate biopolymer synthesis mechanism and improvement of production.

Thermophilic microorganism *Brevibacillus thermoruber* 423 is able to produce high levels of EPS (Yasar Yildiz et al., 2015). Recently, draft genome sequence and whole-genome analysis of this bacterium have been reported. Whole-genome analysis of this bacterium was performed by a systems-based approach to understand the biological mechanisms and whole-genome organization of thermophilic EPS producers. Therefore, strategies for the genetic and metabolic optimization of EPS production could be developed. Genome annotation was used to detect essential genes associated with EPS biosynthesis and a hypothetical mechanism for EPS biosynthesis was generated considering the experimental evidences. The genome sequence of *B. thermoruber* strain 423 is being used to reconstruct a genome-scale metabolic model to develop metabolic engineering strategies since the metabolic model will be used to optimize medium compositions, to modify the microorganism genetically, to improve production yields, and to modify EPS monomer composition (Yasar Yildiz et al., 2014).

Nadkarni et al. (2014) performed comparative genome analysis of *Lactobacillus rhamnosus* clinical isolates to identify EPS cluster. In this study, transcriptional orientation of the *eps* cluster genes, the presence of two genes homologous to priming glycosyltransferases, the absence of *rmlACBD* genes involved in the dTDP-rhamnose biosynthetic pathway, and the presence of a family 2 GTF in the *eps* cluster of both clinical isolates of *L. rhamnosus*, is predicted to alter EPS composition and could influence pathogenicity.

The first complete genome sequence of Russia origin *Bifidobacterium longum* subsp. *longum* strain GT15, comparative genome analysis, identification, and characterization of regulatory genes, *in silico* analysis of all the most significant probiotic genes and considered genes have been reported. The genomic analysis for polysaccharides was also performed, and it was observed that most of the genes in the carbohydrate metabolism category were involved in the utilization of oligo-polysaccharides. The genome also contains genes predicted to encode proteins involved in the production of capsular EPS, which are most likely involved in bacteria–host interactions (Zakharevich et al., 2015).

Genome sequence of moderately halophilic and EPS-producing *Salipiger mucosus* DSM 16094T and the presence of a high number of genes associated with biosynthesis of EPSs have been reported. Genes associated with the synthesis of polyhydroxyalkanoates have been also found (Riedel et al., 2014).

*Acidithiobacillus ferrooxidans* was the first biomining microorganism whose genome was sequenced and the genes involved in the biosynthesis EPSs precursors have been studied (Valdés et al., 2008). The cluster of five genes proposed to be involved in the biosynthesis of EPSs precursors via the Leloir pathway have been also identified previously (Barreto et al., 2005).

A curdland producer *Agrobacterium* sp. ATCC 31749's genome was sequenced and the curdland biosynthesis operon (*crdASC*) was identified (Ruffing et al., 2011). Moreover, transcriptome analysis of this microorganism has been performed to understand the regulation of EPS biosynthesis (Ruffing and Chen, 2012). In this study, transcriptome profiling was used to identify genes that expressed during curdland biosynthesis and carry out targeted gene knockouts to investigate their roles in the transcriptional



regulation of curdian production. The analysis showed that curdian synthesis operon was upregulated by up to 100-fold upon nitrogen depletion. Moreover, novel regulation mechanisms including RpoN-independent NtrC regulation and intracellular pH regulation by acidocalcisomes were identified.

Engineering studies for LAB, particularly members of the genera *Lactococcus* and *Lactobacillus*, have been performed for the production of platform chemicals, such as lactate L- and D-stereoisomers, 1,3-propanediol, and 2,3-butanediol, food flavors and sweeteners, vitamins, and complex polysaccharides (Gaspar et al., 2013).

Exopolysaccharide biosynthetic pathways in LAB have been engineered to have challenges greater than manipulating specific steps in primary metabolism (Patel et al., 2011). Since EPSs enhance potential health benefits of fermented food products, many metabolic engineering approaches are employed to improve productivity and structure of EPS (Gaspar et al., 2013).

Omics studies can be used to improve the understanding of metabolism in food industry microorganisms. The metagenomic studies for fermented food were performed to analyze the metabolic potentials of LAB bacteria, which is very important in industrial fermentations. Jung et al. (2011) performed metagenomic studies changes in bacterial populations, metabolic potential, and the overall genetic features of the microbial community during a 29-day fermentation process of the traditional Korean food kimchi. The transcriptome response has been analyzed in yogurt fermentation (Sieuwerdt et al., 2010) and in milk (Goh et al., 2011).

The first genome-scale model for *L. lactis* (Oliveira et al., 2005), and since whole-genome metabolic reconstructions for *Lb. plantarum* (Teusink et al., 2005) and *S. thermophilus* (Pastink et al., 2009) have been reported. Functional genomics and other studies have performed to investigate genomic diversity in LAB and the findings highlighted the variety of carbon substrates potentially used by LAB including simple sugars, complex carbohydrates such as xylan, starch, and fructans,  $\alpha$ -galactosides (e.g., raffinose and stachyose), pentoses (D arabinose and D-xylose), and the cheap C3 carbon source glycerol (Teusink et al., 2009; Siezen and van Hylckama Vlieg, 2011).

Complete genomic sequence of *Lb. bulgaricus* 2038 has been reported and genomic analysis of EPS biosynthesis has been performed. Two neighboring *eps* clusters with significant differences were identified when compared with genome sequence of *Lb. bulgaricus* species by comparative genomic analysis (Hao et al., 2011).

Genomic studies microbial EPSs producer of deep-sea bacteria such as *Zunongwangia profunda* SM-A87, *Pseudoalteromonas* sp. SM9913, *Pseudoalteromonas haloplanktis* TAC125 (Qin et al., 2011), have been also performed and analyzed for EPS gene clusters (Finore et al., 2014). In addition, genome sequence of several deep-sea isolates such as *Idiomarina loihiensis* (Hou et al., 2004) and *Alteromonas macleodii* (Ivars-Martinez et al., 2008) demonstrated the EPS biosynthesis genes (Finore et al., 2014).

Such omic studies and works from systems biology perspective will play an important role both scientifically and economically, since there is a great need for developing efficient methodologies for enhanced EPS biosynthesis. More information on the genome of the microorganism will enable to develop strategies

to successfully enhance production rate and also to engineer EPSs properties by modifying composition and chain length. Systems-based modeling approach constitutes an important step toward understanding the interplay between metabolism and EPS biosynthesis.

Systems biology approaches and metabolic reconstruction studies of microbial EPS have been also reported for two important EPSs: xanthan and levan. These studies and their findings were given in detail in the following sections. Besides, these two important EPSs researches on pullulan and dextran were also discussed and the general properties and applications for all these important EPSs were summarized in **Table 1**.

## XANTHAN

The commercially most important microbial EPS is known as Xanthan which is produced by the plant-pathogen-proteobacterium *Xanthomonas campestris* pv. *campestris* (Xcc) (Vorhölter et al., 2008; Frese et al., 2014). It is a heteropolysaccharide composed of repetitive pentasaccharide units consisting of monomeric units of two glucose, two mannose, and one GlcA residues with a backbone chain consisting of (1-4)- $\beta$ -D-glucan cellulose (Khan et al., 2007). Due to its superior properties and rheological characteristics, xanthan has found a wide range of applications as a thickening or stabilizing agent in food, cosmetics and oil drilling industries (Schatschneider et al., 2013; Chivero et al., 2015). It has been described as a “benchmark” product based on its significance in food and non-food applications which include dairy products, drinks, confectionary, dressing, bakery products, syrups, and pet foods, as well as the oil, pharmaceutical, cosmetic, paper, paint, and textile industries (Patel and Prajapati, 2013; Cho and Yoo, 2015). Xanthan was also employed in non-food applications such stabilizing cattle feed supplements, calf milk substitutes, agricultural herbicides, fungicides, pesticides, and fertilizers, and to impart thixotropy into toothpaste preparations (Morris, 2006). Considering the commercial importance of this microbial EPS, omics studies and metabolic model reconstructions were performed to clarify xanthan biosynthesis mechanism.

The genomes of five *Xanthomonas* strains *X. campestris* pv. *campestris* strains ATCC 33913 (da Silva et al., 2002) and 8004 (Qian et al., 2005), *X. campestris* pv. *vesicatoria* strain 85-10 (Thieme et al., 2005), *X. oryzae* pv. *oryzae* strains KACC10331 (Lee et al., 2005a,b) and MAFF 311018 (Ochiai et al., 2005), *X. axonopodis* pv. *citri* strain 306 (da Silva et al., 2002) have been sequenced. The complete genome sequence of the xanthan producer strain *Xanthomonas campestris* pv. *campestris* strain B100 and its use for mechanistic model for biosynthesis of xanthan have been also reported (Vorhölter et al., 2008). In this study, the gene products and metabolic pathways for xanthan polymerization were investigated in detail. The gene products of *gumJ*, *gumC*, *gumD*, and *gumE* were analyzed to establish detailed functions in a xanthan polymerization. Moreover, the mechanistic model for the biosynthesis of xanthan was established.

*In vivo* proteome analysis *X. campestris* pv. *campestris* has been performed to investigate protein expression of the microorganism during host–plant interaction. Peptide mass fingerprinting or *de novo* sequencing methods were utilized for identification of the

functions of proteins. This approach will be used to determine the roles of proteins in pathogenicity mechanisms and also xanthan biosynthesis (Andrade et al., 2008).

Schatschneider et al. (2013) has reported the first large-scale metabolic model for *Xanthomonas campestris* (Xcc), which was reconstructed from genome data of, manually curated and further expanded in size. The impact of xanthan production was studied *in vivo* and *in silico* and compared with *gumD* mutant strain. This verified metabolic model is also the first model focusing on bacterial EPS synthesis and it can be used for detailed systems biology analyses and synthetic biology reengineering of Xcc. Moreover, draft genome of *X. campestris* B-1459, which was used in pioneering studies of xanthan biotechnology, and it can be used to analyze the genetic basis of xanthan biosynthesis has been reported recently (Wibberg et al., 2015).

## LEVAN

Levan is a naturally occurring polymer that is composed of  $\beta$ -D-fructofuranose with  $\beta$ (2-6) linkages between fructose rings. It is synthesized by the action of a secreted levansucrase (EC 2.4.1.10) that directly converts sucrose into the polymer (Han and Clarke, 1990). As a homopolysaccharide with many distinguished properties such as high solubility in oil and water, strong adhesivity, good biocompatibility, and film-forming ability, it has great potential as a novel functional biopolymer in foods, feeds, cosmetics, pharmaceutical, and chemical industries (Kang et al., 2009; Kazak Sarilmiser et al., 2015). In fact, a recent literature analysis on microbial EPSs attributed levan together with xanthan, curdlan, and pullulan as the most promising polysaccharides for various industrial sectors (Donot et al., 2012).

Due to its exceptionally high production costs, levan could never find its proper place in the polymer market, and therefore, high-level levan producing microbial systems gain escalating industrial importance. Levan is produced as an EPS from sucrose-based substrates by a variety of microorganisms, including the halophilic bacterium *Halomonas smyrnensis* AAD6<sup>T</sup>, which has been reported as the first levan producer extremophile (Poli et al., 2009).

The gram-negative halophilic bacterium *H. smyrnensis* AAD6<sup>T</sup>, which was isolated from Çamaltı Saltern Area in Turkey (Poli et al., 2009, 2013), was found to excrete high levels of levan

(Poli et al., 2009). With this microbial system, productivity levels were further improved by use of cheap sucrose substitutes such as molasses (Kucukasik, 2010; Kucukasik et al., 2011) as well as other cheap biomass resources (Toksoy Oner, 2013) as fermentation substrate. Further research on the potential uses of levan produced by *H. smyrnensis* AAD6<sup>T</sup> as a bioflocculating agent (Sam et al., 2011), its nanostructured thin films (Sima et al., 2011, 2012), its suitability for peptide-based drug nanocarrier systems (Sezer et al., 2011), and its adhesive multilayer films (Costa et al., 2013) have been reported.

Increasing significance of levan in industrial and medical biotechnology calls for the elucidation of the interrelations between metabolic pathways and levan biosynthesis mechanism in order to control and hence enhance its microbial productivity. However, there is very limited information about the mechanisms involved in the biosynthesis of levan from extremophiles (Nicolaus et al., 2010) and no report about a systematic approach to analyze levan production by *H. smyrnensis* AAD6<sup>T</sup>. Considering this fact, systems-based approaches were applied to improve the levan production capacity of *H. smyrnensis* AAD6<sup>T</sup> cultures (Ates et al., 2013).

The genome sequence forms the basis for metabolic model reconstruction; however, there was a lack for genome information of *H. smyrnensis* AAD6<sup>T</sup>. Only recently, its draft genome sequence has been announced. *De novo* assembly of the whole sequencing reads were carried out in this study. Consequently, several genes related to EPS biosynthesis, including the genes for levansucrase and *ExoD* were revealed by genome analysis (Sogutcu et al., 2012). Due to the absence of genomic information, first, comprehensive metabolic model of a taxonomically close halophilic bacterium, namely, *C. salexigens* DSM3043 have been reconstructed (Ates et al., 2011). Then, in order to investigate levan biosynthesis by a metabolic systems analysis approach, the genome-scale metabolic network of *C. salexigens* was recruited and adopted to *H. smyrnensis* AAD6<sup>T</sup> via integration of the available biochemical, physiological, and phenotypic features of *H. smyrnensis* AAD6<sup>T</sup>. The *in silico* metabolic model was verified with dynamic experimental data on different medium compositions and was then systematically analyzed to identify critical network elements (i.e., enzymes, biochemical transformations, and metabolites) related to levan biosynthesis mechanism. The findings manifested mannitol as a significant metabolite for levan biosynthesis, which was further

**TABLE 1 | General description of significantly important microbial EPSs.**

Microbial EPS	Monomeric units	Microorganism	Industrial applications	Omic studies	Metabolic model
Xanthan	Glucose, mannose and glucuronate	<i>Xanthomonas</i> sp.	Thickening, stabilizing agent, food additive, etc.	Genome sequence, Proteomics	Available (Schatschneider et al., 2013)
Levan	Fructose	<i>Halomonas smyrnensis</i> AAD6 <sup>T</sup> , <i>Zymomonas mobilis</i> , <i>Bacillus subtilis</i> (natto)	Emulsifier, stabilizer and thickener, encapsulating agent, food and feed additive, osmoregulator, and cryoprotector, etc.	Genome sequence	Available (Ates et al., 2013; Diken et al., 2015)
Pullulan	Glucose	<i>Aureobasidium pullulans</i>	Thickening, stabilizing, texturizing, gelling agents, etc.	Genome sequence, proteomics, genome shuffling	Not available
Dextran	Glucose	<i>Leuconostoc</i> spp. and <i>Streptococcus</i> spp.	Blood plasma extender and chromatography media	Genome sequence	Not available

verified experimentally. In the previous study, 1.844 g/L levan yield from the stationary phase bioreactor cultures using a defined media containing sucrose as sole carbon source with almost four-fold increase of levan production was reported (Ates et al., 2013).

Recently, Diken et al. (2015) performed the whole-genome analysis of *H. smyrnensis* AAD6<sup>r</sup> to investigate biological mechanisms, and furthermore, the genome-scale metabolic model iKYA1142, which included 980 metabolites and 1142 metabolic reactions, was reconstructed. The genomic analysis figured out the biotechnological potential of this microorganism as a result of its capacity to produce levan, Pel exopolysaccharide, polyhydroxyalkanoates (PHA), and osmoprotectants. Genes related to EPS biosynthesis and intracellular PHA biosynthesis were detected. *Hs\_SacB* gene encoding the extracellular levansucrase enzyme (EC 2.4.1.10), Pel polysaccharide gene cluster (*PelA*, *PelB*, *PelC*, *PelD*, *PelE*, *PelF*, and *PelG*), Alginate lyase precursor (EC 4.2.2.3), and “Alginate biosynthesis protein Alg8” genes were predicted. The genome information and metabolic model will have a significant role on levan biosynthesis since they will be utilized to improve levan production by metabolic engineering strategies and medium optimization.

## PULLULAN

A fungal glucan “pullulan” that is a linear homopolysaccharide composed of maltotriose reduplicative units connected by  $\alpha$ -1,4-linkages is produced by *Aureobasidium pullulans* (Singh et al., 2008; Sheng et al., 2014; Özcan et al., 2014). Although pullulan production studies mostly focused on *A. pullulans*, other producer strains such as *Remella mesenterica*, *Cryphonectria parasitica*, and *Teloschistes flavicans* were also reported (Cheng et al., 2011). Biosynthesis of pullulan was occurred intracellularly at the cell wall or cell membrane and microorganisms secreted this EPS out to the cell surface to form a loose and slimy layer (Ma et al., 2015). Its solubility in water is excellent as a result of the linkage pattern. Moreover, it has outstanding chemical and physical properties, such as low viscosity, non-toxicity, slow digestibility, high plasticity, and excellent film-forming (Cheng et al., 2011; Sheng et al., 2014).

The major market for pullulan is food industry and also its potential applications in pharmaceutical, agricultural, chemical, cosmetic, biomedical, and environmental remediation areas have been reported (Özcan et al., 2014; Ma et al., 2015). Due to its superior properties, pullulan can be used as non-polluting wrapping material for food supplements, oxygen impermeable, edible, and biodegradable and highly water soluble films, denture adhesives, emulsifiers and stabilizers for various food products, binder, lubricant, gelling agent, oral care products, and blood plasma substitute. Since it is non-immunogenic, non-toxic, non-carcinogenic, and non-mutagenic, pullulan can find novel application areas such as gene therapy, targeting drugs, and gene delivery (Cheng et al., 2011; Singh et al., 2015). Pullulan has been commercialized in various countries, used as a safe food ingredient and pharmaceutical bulking agent and in Japan.

The molecular biosynthesis of pullulan is a complex metabolic process therefore the molecular basis of this process is not clearly identified. The three significant enzymes for pullulan biosynthesis

are  $\alpha$ -phosphoglucose mutase (PGM, EC 5.4.2.2), UDP-glucose pyrophosphorylase (UGP, EC 2.7.7.9), and glucosyltransferase (FKS, EC 2.4.1.34). Nitrogen is the major component for the cultivation of *A. pullulans* in pullulan fermentation (Wang et al., 2015). Wang et al. (2015) analyzed gene expression of the key enzymes (PGM, UGP, FKS) to improve pullulan production. They have reported the improved pullulan production correlated to the high activities of PGM and FKS, increased the activities of  $\alpha$ -phosphoglucose mutase and glucosyltransferase, and upregulated the transcriptional levels of *pgm1* and *fks* genes by nitrogen limitation.

Kang et al. (2011) studied genome shuffling *A. pullulans* N3.387 by ethyl methane sulfonate (EMS) and ultraviolet (UV) mutagenesis to improve pullulan biosynthesis and developed a mutant that could produce more pullulan than wild type strain.

Sheng et al. (2014) performed proteomic studies of pullulan production in *A. pullulans* to understand the effect of different concentrations of  $(\text{NH}_4)_2\text{SO}_4$  which would be useful to optimize industrial pullulan production. The proteomic studies demonstrated the expression of antioxidant related enzymes and energy-generating enzymes and the depression of the enzymes concerning amino acid biosynthesis, glycogen biosynthesis, glycolysis, protein transport, and transcriptional regulation under nitrogen limitation, resulted in conversion of metabolic flux from the glycolysis pathway to the pullulan biosynthesis pathway.

Draft genome of *Aureobasidium pullulans* AY4 was determined and genome analysis revealed the presences of genes coding for commercially important enzymes such as pullulanases, dextranases, amylases, and cellulases (Chan et al., 2012). Gostin et al. (2014) performed *de novo* genome sequencing and genome analysis of the four varieties of *A. pullulans* to investigate the genomic basis of pullulan biosynthesis potential. Single-copy genes for phosphoglucose mutase and uridine diphosphoglucose pyrophosphorylase, which are the key enzymes for converting glucose units into pullulan, were present in four varieties of *A. pullulans*. Genomic analysis also revealed that these microorganisms included all of the putative enzymes that were proposed to be involved in pullulan biosynthesis (Gostin et al., 2014). The genome sequences of *A. pullulans* species will be facilitated to clarify biosynthesis mechanism of pullulan and improve its production.

## DEXTRAN

Dextran, which is a homopolysaccharides composed of  $\alpha$ -1,6 glycosidic linkages, is produced by *L. mesenteroides* hydrolases in the presence of sucrose. Various strains of bacteria such as *Streptococcus* and *Acetobacter* have been also found to produce dextran bacteria (e.g., *Leuconostoc* and *Streptococcus*), through the use of specific enzymes like glucansucrases (Patel et al., 2011; Casettari et al., 2015). Since dextran has a flexible structure as a result of free rotation of glycosidic bond and it is highly soluble in water, biocompatible, and biodegradable, it becomes a functional hydrocolloid (Ahmad et al., 2015). It has commercial applications in food, pharmaceutical and chemical industries as adjuvant, emulsifier, thickener, carrier, and stabilizer. Moreover, dextran is used as therapeutic agent to restore blood volume, for the matrix preparation of chromatography



columns, for synthesizing dextran sulfate for blood coagulation prevention and blood flow facilitation, as osmotic agents, lubricant in eye drops and for increase of blood sugar levels, iron carrier, anticoagulant for pharmacy (Patel et al., 2011; Han et al., 2014).

Dextran has also found use in different areas such as veterinary medicines, biosensor material, food syrup stabilizers, dough improvers, metal-plating processes, enhanced oil recovery, stabilizing coating for protecting metal nanoparticles against oxidation, and coating on biomaterials to prevent undesirable protein absorption. Dextran and its derivatives such as cyclodextran are utilized in the pharmaceutical industry like cariostatic, anti-HIV, and anti-ulcer agent (Ahmad et al., 2015; Casettari et al., 2015).

Various genome sequence and genome analysis studies were carried out with dextran producer strains. For instance, the genome of LAB *Leuconostoc gasicomitatum* was sequenced and genome analysis was performed to understand the growth and spoilage potentials. The genomic analysis revealed genes for two dextransucrases catalyzing the formation of dextran from sucrose: *epsA* (LEGAS\_0699) is part of a large EPS cluster, while *dsrA* (LEGAS\_1012) is located as a single gene in the chromosome (Johansson et al., 2011). Saulnier et al. (2011) performed the metabolic pathway reconstruction, genome profiling, genomic, and transcriptomic comparisons of the *Lactobacillus reuteri* strains to define functional probiotic features. Genome-wide comparison resulted in dextranase gene that predicted to encode the synthesis of EPS. The first genome sequence of *Weissella* species has been announced for *Weissella confusa* (formerly *Lactobacillus confusus*) LBAE C39-2, which was also found promising for *in situ* production of dextran in sourdoughs (Amari et al., 2012). A draft genome sequence of dextran producer *Leuconostoc lactis* EFEL005 has been announced and genomic analysis was performed to understand its probiotic properties as a starter for fermented foods (Moon et al., 2015). The increased number of

genome sequences will accelerate systems-based works for dextran biosynthesis mechanism.

## CONCLUSION AND RECOMMENDATIONS

In microbial EPS production, a better understanding of biosynthesis mechanism is a significant issue for optimization of production yields, improvement of product quality and properties, and also for the design of novel strains. As most of the novel bacterial EPS with unique properties have expensive production costs and economic hurdles need to be overcome, this valuable information about biosynthesis is also important to lower these charges.

More information on the genome of the EPS producer microorganisms will enable to develop additional strategies to successfully enhance EPSs production rate and also to engineer their properties by modifying composition and chain length. Since genome-scale reconstruction includes every reaction of target organism through integrating genome annotation and biochemical information, a systems-based metabolic modeling approach constitutes an important step toward understanding the interplay between metabolism and EPSs biosynthesis. Since microbial biopolymer biosynthesis is a result of a complex system of many metabolic processes, systems-based approaches are needed to control and optimize production in order to improve the formerly reported yields.

Furthermore, the genome-scale metabolic model based on genome sequence will have the capacity to consider gene expression, metabolomics, and proteomics data to get accurate prediction at different environmental conditions.

## ACKNOWLEDGMENTS

The financial support provided by TUBITAK through project 114M239 is gratefully acknowledged.

## REFERENCES

- Ahmad, N. H., Mustafa, S., and Che Man, Y. B. (2015). Microbial polysaccharides and their modification approaches: a review. *Int. J. Food Prop.* 18, 332–347. doi:10.1080/10942912.2012.693561
- Aldor, I. S., and Keasling, J. D. (2003). Process design for microbial plastic factories: metabolic engineering of polyhydroxyalkanoates. *Curr. Opin. Biotechnol.* 14, 475–483. doi:10.1016/j.copbio.2003.09.002
- Amari, M., Laguerre, S., Vuillemin, M., Robert, H., Loux, V., Klopp, C., et al. (2012). Genome sequence of *Weissella confusa* LBAE C39-2, isolated from a wheat sourdough. *J. Bacteriol.* 194, 1608–1609. doi:10.1128/JB.06788-11
- Andrade, A. E., Silva, L. P., Pereira, J. L., Noronha, E. F., Reis, F. B. Jr., Bloch, C. Jr., et al. (2008). In vivo proteome analysis of *Xanthomonas campestris* pv. *campestris* in the interaction with the host plant *Brassica oleracea*. *FEMS Microbiol. Lett.* 281, 167–174. doi:10.1111/j.1574-6968.2008.01090.x
- Arco, Y., Llamas, I., Martinez-Checa, F., Argandona, M., Quesada, E., and del Moral, A. (2005). *epsABCJ* genes are involved in the biosynthesis of the exopolysaccharide mauran produced by *Halomonas maura*. *Microbiology* 151, 2841–2851. doi:10.1099/mic.0.27981-0
- Ates, O., Arga, K. Y., and Toksoy Öner, E. (2013). The stimulatory effect of mannitol on levan biosynthesis: lessons from metabolic systems analysis of *Halomonas smyrnensis* AAD6<sup>T</sup>. *Biotechnol. Prog.* 29, 1386–1397. doi:10.1002/btpr.1823
- Ates, O., Toksoy Öner, E., and Arga, K. Y. (2011). Genome-scale reconstruction of metabolic network for a halophilic extremophile, *Chromohalobacter salexigens* DSM 3043. *BMC Syst. Biol.* 5:12. doi:10.1186/1752-0509-5-12
- Baart, G. J., and Martens, D. E. (2012). “Genome-scale metabolic models: reconstruction and analysis,” in Christodoulides M. editor. *Neisseria Meningitidis* (London: Humana Press), 107–126.
- Badel, S., Bernardi, T., and Michaud, P. (2011). New perspectives for *Lactobacilli* exopolysaccharides. *Biotechnol. Adv.* 29, 54–66. doi:10.1016/j.biotechadv.2010.08.011
- Barreto, M., Jedlicki, E., and Holmes, D. S. (2005). Identification of a gene cluster for the formation of extracellular polysaccharide precursors in the chemolithoautotroph *Acidithiobacillus ferrooxidans*. *Appl. Environ. Microbiol.* 71, 2902–2909. doi:10.1128/AEM.71.6.2902-2909.2005
- Barrett, C. L., Kim, T. Y., Kim, H. U., Palsson, B., and Lee, S. Y. (2006). Systems biology as a foundation for genome-scale synthetic biology. *Curr. Opin. Biotechnol.* 17, 488–492. doi:10.1016/j.copbio.2006.08.001
- Becker, A. (2015). Challenges and perspectives in combinatorial assembly of novel exopolysaccharide biosynthesis pathways. *Front. Microbiol.* 6:687. doi:10.3389/fmicb.2015.00687
- Boels, I. C., Kranenburg, R. V., Hugenholtz, J., Kleerebezem, M., and de Vos, W. M. (2001). Sugar catabolism and its impact on the biosynthesis and engineering of exopolysaccharide production in lactic acid bacteria. *Int. Dairy J.* 11, 723–732. doi:10.1016/S0958-6946(01)00116-9
- Cai, L., Tan, D., Aibaidula, G., Dong, X., Chen, J. C., Tian, W. D., et al. (2011). Comparative genomics study of polyhydroxyalkanoates (PHA) and ectoine relevant genes from *Halomonas* sp. TD01 revealed extensive horizontal gene transfer events and co-evolutionary relationships. *Microb. Cell Fact.* 10, 88.



- Casettari, L., Bonacucina, G., Morris, G. A., Perinelli, D. R., Lucaioli, P., Cespi, M., et al. (2015). Dextran and its potential use as tablet excipient. *Powder Technol.* 273, 125–132. doi:10.1016/j.powtec.2014.12.030
- Cescutti, P., Foschiatti, M., Furlanis, L., Lagatolla, C., and Rizzo, R. (2010). Isolation and characterisation of the biological repeating unit of cepacian, the exopolysaccharide produced by bacteria of the *Burkholderia cepacia* complex. *Carbohydr. Res.* 345, 1455–1460. doi:10.1016/j.carres.2010.03.029
- Chan, G. F., Bamadhaj, H. M., Gan, H. M., and Rashid, N. A. A. (2012). Genome sequence of *Aureobasidium pullulans* AY4, an emerging opportunistic fungal pathogen with diverse biotechnological potential. *Eukaryotic Cell* 11, 1419–1420. doi:10.1128/EC.00245-12
- Cheng, K. C., Demirci, A., and Catchmark, J. M. (2011). Pullulan: biosynthesis, production, and applications. *Appl. Microbiol. Biotechnol.* 92, 29–44. doi:10.1007/s00253-011-3477-y
- Chivero, P., Gohtani, S., Yoshii, H., and Nakamura, A. (2015). Effect of xanthan and guar gums on the formation and stability of soy soluble polysaccharide oil-in-water emulsions. *Food Res. Intern.* 70, 7–14. doi:10.1016/j.foodres.2015.01.025
- Cho, H. M., and Yoo, B. (2015). Rheological characteristics of cold thickened beverages containing xanthan gum-based food thickeners used for dysphagia diets. *J. Acad. Nutr. Diet.* 115, 106–111. doi:10.1016/j.jand.2014.08.028
- Costa, R. R., Neto, A. I., Calgeris, I., Correia, C. R., Pinho, A. C., Fonseca, J., et al. (2013). Adhesive nanostructured multilayer films using a bacterial exopolysaccharide for biomedical applications. *J. Mater. Chem. B* 1(18), 2367–2374.
- Crescenzi, V., and Dentini, M. (1996). "Microbes in polymer chemistry," in *Hydrogels and Biodegradable Polymers for Bioapplications*, ACS Symposium Series, Vol. 627, ed. Ottenbrite R. M., Huang S. J., and Park K. (Washington, DC: American Chemical Society), 221–232.
- Curran, K. A., and Alper, H. S. (2012). Expanding the chemical palate of cells by combining systems biology and metabolic engineering. *Metab. Eng.* 14, 289–297. doi:10.1016/j.jmben.2012.04.006
- da Silva, A. R., Ferro, J. A., Reinach, F. C., Farah, C. S., Furlan, L. R., Quaggio, R. B., et al. (2002). Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* 417, 459–463. doi:10.1038/417459a
- Delbarre-Ladrat, C., Sinquin, C., Lebellenger, L., Zykwinska, A., and Collic-Jouault, S. (2014). Exopolysaccharides produced by marine bacteria and their applications as glycosaminoglycan-like molecules. *Front. Chem.* 2:85. doi:10.3389/fchem.2014.00085
- Diken, E., Ozer, T., Arikian, M., Emrence, Z., Oner, E. T., Ustek, D., et al. (2015). Genomic analysis reveals the biotechnological and industrial potential of levan producing halophilic extremophile, *Halomonas smyrnensis* AAD6<sup>T</sup>. *Springerplus* 4, 393. doi:10.1186/s40064-015-1184-3
- Donot, F., Fontana, A., Baccou, J. C., and Schorr-Galindo, S. (2012). Microbial exopolysaccharides: main examples of synthesis, excretion, genetics and extraction. *Carbohydr. Polym.* 87, 951–962. doi:10.1016/j.carbpol.2011.08.083
- Esvelt, K. M., and Wang, H. H. (2013). Genome-scale engineering for systems and synthetic biology. *Mol. Syst. Biol.* 9, 641. doi:10.1038/msb.2012.66
- Finore, I., Di Donato, P., Mastascusa, V., Nicolaus, B., and Poli, A. (2014). Fermentation technologies for the optimization of marine microbial exopolysaccharide production. *Mar. Drugs* 12, 3005–3024. doi:10.3390/md12053005
- Flemming, H. C., and Wingender, J. (2010). The biofilm matrix. *Nat. Rev. Microbiol.* 8, 623–633. doi:10.1038/nrmicro2415
- Freitas, F., Alves, V. D., and Reis, M. A. (2011). Advances in bacterial exopolysaccharides: from production to biotechnological applications. *Trends Biotechnol.* 29, 388–398. doi:10.1016/j.tibtech.2011.03.008
- Frese, M., Schatschneider, S., Voss, J., Vorhölder, F. J., and Niehaus, K. (2014). Characterization of the pyrophosphate-dependent 6-phosphofructokinase from *Xanthomonas campestris* pv. campestris. *Arch. Biochem. Biophys.* 546, 53–63. doi:10.1016/j.abb.2014.01.023
- Gaspar, P., Carvalho, A. L., Vinga, S., Santos, H., and Neves, A. R. (2013). From physiology to systems metabolic engineering for the production of biochemicals by lactic acid bacteria. *Biotechnol. Adv.* 31, 764–788. doi:10.1016/j.biotechadv.2013.03.011
- Goh, Y. J., Goin, C., O'Flaherty, S., Altermann, E., and Hutkins, R. (2011). Specialized adaptation of a lactic acid bacterium to the milk environment: the comparative genomics of *Streptococcus thermophilus* LMD-9. *Microb. Cell Fact.* 10(Suppl. 1), S22. doi:10.1186/1475-2859-10-S1-S22
- Gostin, C., Ohm, R. A., Kogej, T., Sonjak, S., Turk, M., Zajc, J., et al. (2014). Genome sequencing of four *Aureobasidium pullulans* varieties: biotechnological potential, stress tolerance, and description of new species. *BMC Genomics* 15:549. doi:10.1186/1471-2164-15-549
- Han, J., Hang, F., Guo, B., Liu, Z., You, C., and Wu, Z. (2014). Dextran synthesized by *Leuconostoc mesenteroides* BD1710 in tomato juice supplemented with sucrose. *Carbohydr. Polym.* 112, 556–562. doi:10.1016/j.carbpol.2014.06.035
- Han, Y. W., and Clarke, M. A. (1990). Production and characterization of microbial levan. *J. Agric. Food Chem.* 38, 393–396. doi:10.1021/jf00092a011
- Hanemaaijer, M., Röling, W. F., Olivier, B. G., Khandelwal, R. A., Teusink, B., and Bruggeman, F. J. (2015). Systems modeling approaches for microbial community studies: from metagenomics to inference of the community structure. *Front. Microbiol.* 6:213. doi:10.3389/fmicb.2015.00213
- Hao, P., Zheng, H., Yu, Y., Ding, G., Gu, W., Chen, S., et al. (2011). Complete sequencing and pan-genomic analysis of *Lactobacillus delbrueckii* subsp. *Bulgarius* reveal its genetic basis for industrial yogurt production. *PLoS ONE* 6:e15964. doi:10.1371/journal.pone.0015964
- Henry, C. S., DeJongh, M., Best, A. A., Frybarger, P. M., Linsay, B., and Stevens, R. L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* 28, 977–982. doi:10.1038/nbt.1672
- Hong, S. H., Park, S. J., Moon, S. Y., Park, J. P., and Lee, S. Y. (2003). In silico prediction and validation of the importance of the Entner-Doudoroff pathway in poly (3-hydroxybutyrate) production by metabolically engineered *Escherichia coli*. *Biotechnol. Bioeng.* 83, 854–863. doi:10.1002/bit.10733
- Hou, S., Saw, J. H., Lee, K. S., Freitas, T. A., Belisle, C., Kavarabaysi, Y. M., et al. (2004). Genome sequence of the deep-sea  $\gamma$ -proteobacterium *Idiomarina loihiensis* reveals amino acid fermentation as a source of carbon and energy. *Proc. Natl. Acad. Sci. U.S.A.* 101, 18036–18041. doi:10.1073/pnas.0407638102
- Ivars-Martinez, E., Martin-Cuadrado, A. B., D'Auria, G., Mira, A., Ferreira, S., Johnson, J., et al. (2008). Comparative genomics of two ecotypes of the marine planktonic copiotroph *Alteromonas macleodii* suggests alternative lifestyles associated with different kinds of particulate organic matter. *ISME J.* 2, 1194–1212. doi:10.1038/ismej.2008.74
- Johansson, P., Paulin, L., Säde, E., Salovuori, N., Alatalo, E. R., Björkroth, K. J., et al. (2011). Genome sequence of a food spoilage lactic acid bacterium, *Leuconostoc gasicomitatum* LMG 18811T, in association with specific spoilage reactions. *Appl. Environ. Microbiol.* 77, 4344–4351. doi:10.1128/AEM.00102-11
- Jung, J. Y., Lee, S. H., Kim, J. M., Park, M. S., Bae, J. W., and Hahn, Y. (2011). Metagenomic analysis of kimchi, a traditional Korean fermented food. *Appl. Environ. Microbiol.* 77, 2264–2274. doi:10.1128/AEM.02157-10
- Kang, J. X., Chen, X. J., Chen, W. R., Li, M. S., Fang, Y., Li, D. S., et al. (2011). Enhanced production of pullulan in *Aureobasidium pullulans* by a new process of genome shuffling. *Process Biochem.* 46, 792–795. doi:10.1016/j.procbio.2010.11.004
- Kang, S. A., Jang, K. H., Seo, J. W., Kim, K. H., Kim, Y. H., Rairakhwada, D., et al. (2009). "Levan: applications and perspectives," in *Microbial Production of Biopolymers and Polymer Precursors: Applications and Perspectives*, ed. Rehm B. (Norfolk: Caister Academic Press, Inc.), 145–161.
- Kaur, V., Bera, M. B., Panesar, P. S., Kumar, H., and Kennedy, J. F. (2014). Welan gum: microbial production, characterization, and applications. *Int. J. Biol. Macromol.* 65, 454–461. doi:10.1016/j.ijbiomac.2014.01.061
- Kazak Sarilimiser, H., Ates, Ö., Özdemir, G., Arga, K. Y., and Toksoy Öner, E. (2015). Effective stimulating factors for microbial levan production by *Halomonas smyrnensis* AAD6<sup>T</sup>. *J. Biosci. Bioeng.* 119, 455–463. doi:10.1016/j.jbiosc.2014.09.019
- Kennedy, C. J., Boyle, P. M., Waks, Z., and Silver, P. A. (2009). Systems-level engineering of nonfermentative metabolism in yeast. *Genetics* 183, 385–397. doi:10.1534/genetics.109.105254
- Keshavarz, T., and Roy, I. (2010). Polyhydroxyalkanoates: bioplastics with a green agenda. *Curr. Opin. Microbiol.* 13, 321–326. doi:10.1016/j.mib.2010.02.006
- Khan, T., Park, J. K., and Kwon, J. H. (2007). Functional biopolymers produced by biochemical technology considering applications in food engineering. *Korean J. Chem. Eng.* 24, 816–826. doi:10.1007/s11814-007-0047-1
- Kind, S., and Wittmann, C. (2011). Bio-based production of the platform chemical 1, 5-diaminopentane. *Appl. Microbiol. Biotechnol.* 91, 1287–1296. doi:10.1007/s00253-011-3457-2
- Kucukaskis, F. (2010). *Evaluation of Different Waste Streams as Fermentation Substrates for Microbial Biopolymer Production*. Ph.D. thesis, Institute of Environmental Sciences, Istanbul, Turkey.

- Kucukasik, F., Kazak, H., Guney, D., Finore, I., Poli, A., Yenigun, O., et al. (2011). Molasses as fermentation substrate for Levan production by *Halomonas* sp. *Appl. Microbiol. Biotechnol.* 89(6), 1729–1740.
- Kumar, A. S., Mody, K., and Jha, B. (2007). Bacterial exopolysaccharides-a perception. *J. Basic Microbiol.* 47(2), 103–117.
- Kumar, A. S., and Mody, K. (2009). “Microbial exopolysaccharides: variety and potential applications,” in *Microbial Production of Biopolymers and Polymer Precursors: Applications and Perspectives*, ed. Rehm B. (Norfolk: Caister Academic Press, Inc.), 229–253.
- Kwon, K. J., Park, K. J., Kim, J. D., Kong, J. Y., and Kong, I. S. (1994). Isolation of two different polysaccharides from halophilic *Zoogloea* sp. *Biotechnol. Lett.* 16, 783–788. doi:10.1007/BF00133954
- Lee, B. M., Park, Y. J., Park, D. S., Kang, H. W., Kim, J. G., Song, E. S., et al. (2005a). The genome sequence of *Xanthomonas oryzae pathovar oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Res.* 33, 577–586. doi:10.1093/nar/gki206
- Lee, S. J., Lee, D. Y., Kim, T. Y., Kim, B. H., Lee, J., and Lee, S. Y. (2005b). Metabolic engineering of *Escherichia coli* for enhanced production of succinic acid, based on genome comparison and *in silico* gene knockout simulation. *Appl. Environ. Microbiol.* 71, 7880–7887. doi:10.1128/AEM.71.12.7880-7887.2005
- Lee, J. W., Kim, H. U., Choi, S., Yi, J., and Lee, S. Y. (2011). Microbial production of building block chemicals and polymers. *Curr. Opin. Biotechnol.* 22, 758–767. doi:10.1016/j.copbio.2011.02.011
- Lee, J. W., Na, D., Park, J. M., Lee, J., Choi, S., and Lee, S. Y. (2012). Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat. Chem. Biol.* 8, 536–546. doi:10.1038/nchembio.970
- Lerman, J. A., Hyduke, D. R., Latif, H., Portnoy, V. A., Lewis, N. E., Orth, J. D., et al. (2012). *In silico* method for modelling metabolism and gene product expression at genome scale. *Nat. Commun.* 3, 929. doi:10.1038/ncomms1928
- Li, J., and Wang, N. (2012). The *gpsX* gene encoding a glycosyltransferase is important for polysaccharide production and required for full virulence in *Xanthomonas citri* subsp. *Citri*. *BMC Microbiol.* 12:31. doi:10.1186/1471-2180-12-31
- Liang, T. W., and Wang, S. L. (2015). Recent advances in exopolysaccharides from *Paenibacillus* spp.: production, isolation, structure, and bioactivities. *Mar. Drugs* 13, 1847–1863. doi:10.3390/md13041847
- Llomas, I., Amjres, H., Mata, J. A., Quesada, E., and Béjar, V. (2012). The potential biotechnological applications of the exopolysaccharide produced by the halophilic bacterium *Halomonas almeriensis*. *Molecules* 17, 7103–7120. doi:10.3390/molecules17067103
- Long, M. R., Ong, W. K., and Reed, J. L. (2015). Computational methods in metabolic engineering for strain design. *Curr. Opin. Biotechnol.* 34, 135–141. doi:10.1016/j.copbio.2014.12.019
- Ma, Z. C., Liu, N. N., Chi, Z., Liu, G. L., and Chi, Z. M. (2015). Genetic modification of the marine-isolated yeast *Aureobasidium melanogenum* P16 for efficient pullulan production from inulin. *Mar. Biotechnol.* 17, 511–522. doi:10.1007/s10126-015-9638-8
- Mantzaris, N. V., Kelley, A. S., Srienc, F., and Daoutidis, P. (2001). Optimal carbon source switching strategy for the production of PHA copolymers. *Am. Inst. Chem. Eng. J.* 47, 727–743. doi:10.1002/aic.690470319
- Mata, J. A., Béjar, V., Llomas, I., Arias, S., Bressollier, P., Tallon, R., et al. (2006). Exopolysaccharides produced by the recently described halophilic bacteria *Halomonas ventosae* and *Halomonas anticariensis*. *Res. Microbiol.* 157, 827–835. doi:10.1016/j.resmic.2006.06.004
- Medina-Cleghorn, D., and Nomura, D. K. (2013). Chemical approaches to study metabolic networks. *Pfugers Arch.* 465, 427–440. doi:10.1007/s00424-012-1201-0
- Moon, J. S., Choi, H. S., Shin, S. Y., Noh, S. J., Jeon, C. O., and Han, N. S. (2015). Genome sequence analysis of potential probiotic strain *Leuconostoc lactis* EFEL005 isolated from kimchi. *J. Microbiol.* 53, 337–342. doi:10.1007/s12275-015-5090-8
- Moon, S. Y., Hong, S. H., Kim, T. Y., and Lee, S. Y. (2008). Metabolic engineering of *Escherichia coli* for the production of malic acid. *Biochem. Eng. J.* 40, 312–320. doi:10.1016/j.bej.2008.01.001
- Morris, G., and Harding, S. E. (2009). “Polysaccharides, microbial,” in *Encyclopedia of Microbiology*, ed. Schaechter M. (Amsterdam: Elsevier), 482–494.
- Morris, V. J. (2006). “Bacterial polysaccharides,” in *Food Polysaccharides and Their Applications*, eds Stephen A. M., Phillips G. O., and Williams P. A. (New York: Taylor Francis), 413–454.
- Nadkarni, M. A., Chen, Z., Wilkins, M. R., and Hunter, N. (2014). Comparative genome analysis of *Lactobacillus rhamnosus* clinical isolates from initial stages of dental pulp infection: identification of a new exopolysaccharide cluster. *PLoS ONE* 9:e90643. doi:10.1371/journal.pone.0090643
- Nakamura, C. E., and Whited, G. M. (2003). Metabolic engineering for the microbial production of 1, 3-propanediol. *Curr. Opin. Biotechnol.* 14, 454–459. doi:10.1016/j.copbio.2003.08.005
- Ng, C. Y., Jung, M. Y., Lee, J., and Oh, M. K. (2012). Production of 2, 3-butanediol in *Saccharomyces cerevisiae* by *in silico* aided metabolic engineering. *Microb. Cell Fact.* 11, 68. doi:10.1186/1475-2859-11-68
- Nicolaus, B., Kambourova, M., and Toksoy Oner, E. (2010). Exopolysaccharides from extremophiles: from fundamentals to biotechnology. *Environ. Technol.* 31(10), 1145–1158.
- Nogales, J., Palsson, B. O., and Thiele, I. (2008). A genome-scale metabolic reconstruction of *Pseudomonas putida* KT2440: iJN746 as a cell factory. *BMC Syst. Biol.* 2:79. doi:10.1186/1752-0509-2-79
- Nwodo, U. U., Green, E., and Okoh, A. I. (2012). Bacterial exopolysaccharides: functionality and prospects. *Int. J. Mol. Sci.* 13, 14002–14015. doi:10.3390/ijms131114002
- Oberhardt, M. A., Puchalka, J., Fryer, K. E., Martins dos Santos, V. A. P., and Papin, J. A. (2008). Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* 190, 2790–2803. doi:10.1128/JB.01583-07
- Ochiai, H., Inoue, Y., Takeya, M., Sasaki, A., and Kaku, H. (2005). Genome sequence of *Xanthomonas oryzae* pv. *Oryzae* suggests contribution of large numbers of effector genes and insertion sequences to its race diversity. *Jpn. Agric. Res. Q.* 39, 275–287. doi:10.6090/jarq.39.275
- Okuda, S., and Yoshizawa, A. C. (2011). ODB: a database for operon organizations, 2011 update. *Nucleic Acids Res.* 39(Suppl. 1), D552–D555. doi:10.1093/nar/gkq1090
- Oliveira, A. P., Nielsen, J., and Forster, J. (2005). Modeling *Lactococcus lactis* using a genome-scale flux model. *BMC Microbiol.* 5:39. doi:10.1186/1471-2180-5-39
- Ordax, M., Marco-Noales, E., López, M. M., and Biosca, E. G. (2010). Exopolysaccharides favor the survival of *Erwinia amylovora* under copper stress through different strategies. *Res. Microbiol.* 161, 549–555. doi:10.1016/j.resmic.2010.05.003
- Otero, J. M., and Nielsen, J. (2010). Industrial systems biology. *Biotechnol. Bioeng.* 105, 439–460. doi:10.1002/bit.22592
- Özcan, E., Sargin, S., and Goksungur, Y. (2014). Comparison of pullulan production performances of air-lift and bubble column bioreactors and optimization of process parameters in air-lift bioreactor. *Biochem. Eng. J.* 92, 9–15. doi:10.1016/j.bej.2014.05.017
- Pasteur, L. (1861). On the viscous fermentation and the butyrous fermentation. *Bulletin de la Société Chimique de France* 11, 30–31; (in French).
- Pastink, M. I., Teusink, B., Hols, P., Visser, S., de Vos, W. M., and Hugenholtz, J. (2009). Genome-scale model of *Streptococcus thermophilus* LMG18311 for metabolic comparison of lactic acid bacteria. *Appl. Environ. Microbiol.* 75, 3627–3633. doi:10.1128/AEM.00138-09
- Patel, A., and Prajapati, J. B. (2013). Food and health applications of exopolysaccharides produced by lactic acid bacteria. *Adv. Dairy Res.* 1, 2. doi:10.4172/2329-888X.1000107
- Patel, S., Majumder, A., and Goyal, A. (2011). Potentials of exopolysaccharides from lactic acid bacteria. *Indian J. Microbiol.* 52, 3–12. doi:10.1007/s12088-011-0148-8
- Péant, B., LaPointe, G., Gilbert, C., Atlan, D., Ward, P., and Roy, D. (2005). Comparative analysis of the exopolysaccharide biosynthesis gene clusters from four strains of *Lactobacillus rhamnosus*. *Microbiology* 151, 1839–1851. doi:10.1099/mic.0.27852-0
- Poli, A., Kazak, H., Gürlendag, B., Tommonaro, G., Pieretti, G., Toksoy Öner, E., et al. (2009). High level synthesis of levan by a novel *Halomonas* species growing on defined media. *Carbohydr. Polym.* 78, 651–657. doi:10.1016/j.carbpol.2009.05.031
- Poli, A., Nicolaus, B., Denizci, A. A., Yavuzturk, B., and Kazan, D. (2013). *Halomonas smyrnensis* sp. nov., a moderately halophilic, exopolysaccharide-producing bacterium from camalti saltern area, Turkey. *Int. J. Syst. Evol. Microbiol.* 63, 10–18. doi:10.1099/ijse.0.037036-0
- Puchalka, J., Oberhardt, M. A., Godinho, M., Bielecka, A., Regenhart, D., Timmis, K. N., et al. (2008). Genome-scale reconstruction and analysis of the

- Pseudomonas putida* KT2440 metabolic network facilitates applications in biotechnology. *PLoS Comput. Biol.* 4:e1000210. doi:10.1371/journal.pcbi.1000210
- Qian, W., Jia, Y., Ren, S. X., He, Y. Q., Feng, J. X., Lu, L. F., et al. (2005). Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. *Campestris*. *Genome Res.* 15, 757–767. doi:10.1101/gr.3378705
- Qin, Q. L., Li, Y., Zhang, Y. J., Zhou, Z. M., Zhang, W. X., Chen, X. L. Z., et al. (2011). Comparative genomics reveals a deep-sea sediment-adapted life style of *Pseudoalteromonas* sp. SM9913. *ISME J.* 5, 274–284. doi:10.1038/ismej.2010.103
- Rehm, B. (2009). *Microbial Production of Biopolymers and Polymer Precursors: Applications and Perspectives*. Norfolk: Caister Academic.
- Rehm, B. H. (2010). Bacterial polymers: biosynthesis, modifications and applications. *Nat. Rev. Microbiol.* 8, 578–592. doi:10.1038/nrmicro2354
- Riedel, T., Spring, S., Fiebig, A., Petersen, J., Kyrpides, N. C., Göker, M., et al. (2014). Genome sequence of the exopolysaccharide-producing *Salipiger mucosus* type strain (DSM 16094T), a moderately halophilic member of the *Roseobacter clade*. *Stand. Genomic Sci.* 9, 1333. doi:10.4056/sigs.4909790
- Ruffing, A. M., Castro-Melchor, M., Hu, W. S., and Chen, R. R. (2011). Genome sequence of the curdlan-producing *Agrobacterium* sp. strain ATCC 31749. *J. Bacteriol.* 193(16), 4294–4295.
- Ruffing, A. M., and Chen, R. R. (2012). Transcriptome profiling of a curdlan-producing *Agrobacterium* reveals conserved regulatory mechanisms of exopolysaccharide biosynthesis. *Microb. Cell Fact.* 11, 1–13. doi:10.1186/1475-2859-11-17
- Rühmann, B., Schmid, J., and Sieber, V. (2015). High throughput exopolysaccharide screening platform: from strain cultivation to monosaccharide composition and carbohydrate fingerprinting in one day. *Carbohydr. Polym.* 122, 212–220. doi:10.1016/j.carbpol.2014.12.021
- Sam, S., Kucukasik, F., Yenigun, O., Nicolaus, B., Toksoy Oner, E., and Yukselen, M. A. (2011). Flocculating performances of exopolysaccharides produced by a halophilic bacterial strain cultivated on agro-industrial waste. *Bioresour. Technol.* 102, 1788–1794. doi:10.1016/j.biortech.2010.09.020
- Sanchez-Garcia, M. D., Lagaron, J. M., and Hoa, S. V. (2010). Effect of addition of carbon nanofibers and carbon nanotubes on properties of thermoplastic biopolymers. *Compos. Sci. Technol.* 70, 1095–1105. doi:10.1016/j.compscitech.2010.02.015
- Saulnier, D. M., Santos, F., Roos, S., Mistretta, T. A., Spinler, J. K., Molenaar, D., et al. (2011). Exploring metabolic pathway reconstruction and genome-wide expression profiling in *Lactobacillus reuteri* to define functional probiotic features. *PLoS ONE* 6:e18783. doi:10.1371/journal.pone.0018783
- Schatschneider, S., Persicke, M., Watt, S. A., Hublik, G., Pühler, A., Niehaus, K., et al. (2013). Establishment *in silico* analysis, and experimental verification of a large-scale metabolic network of the xanthan producing *Xanthomonas campestris* pv. *Campestris* strain B100. *J. Biotechnol.* 167, 123–134. doi:10.1016/j.jbiotec.2013.01.023
- Schmid, J., and Sieber, V. (2015). Enzymatic transformations involved in the biosynthesis of microbial exopolysaccharides based on the assembly of repeat units. *ChemBiochem* 16, 1141–1147. doi:10.1002/cbic.201500035
- Sezer, A. D., Kazak, H., Toksoy Oner, E., and Akbuğa, J. (2011). Levan-based nanocarrier system for peptide and protein drug delivery: optimization and influence of experimental parameters on the nanoparticle characteristic. *Carbohydr. Polym.* 84, 358–363. doi:10.1016/j.carbpol.2010.11.046
- Sheng, L., Zhu, G., and Tong, Q. (2014). Comparative proteomic analysis of *Aureobasidium pullulans* in the presence of high and low levels of nitrogen source. *J. Agric. Food Chem.* 62, 10529–10534. doi:10.1021/jf503390f
- Sieuwerds, S., Molenaar, D., van Hijum, S. A., Beerthuyzen, M., Stevens, M. J., and Janssen, P. W. (2010). Mixed-culture transcriptome analysis reveals the molecular basis of mixed-culture growth in *Streptococcus thermophilus* and *Lactobacillus bulgaricus*. *Appl. Environ. Microbiol.* 76, 7775–7784. doi:10.1128/AEM.01122-10
- Siezen, R. J., and van Hylckama Vlieg, J. E. (2011). Genomic diversity and versatility of *Lactobacillus plantarum*, a natural metabolic engineer. *Microb. Cell Fact.* 10(Suppl. 1), S3. doi:10.1186/1475-2859-10-S1-S3
- Sima, F., Axente, E., Sima, L. E., Tuyel, U., Eroglu, M. S., Serban, N., et al. (2012). Combinatorial matrix-assisted pulsed laser evaporation: single-step synthesis of biopolymer compositional gradient thin film assemblies. *Appl. Phys. Lett.* 101, 233705. doi:10.1063/1.4769987
- Sima, F., Mutlu Cansever, E., Eroglu, M. S., Sima Livia, E., Serban, N., Ristoscu, C., et al. (2011). Levan nanostructured thin films by MAPLE assembling. *Biomacromolecules* 12, 2251–2256. doi:10.1021/bm200340b
- Simeonidis, E., and Price, N. D. (2015). Genome-scale modeling for metabolic engineering. *J. Ind. Microbiol. Biotechnol.* 42, 327–338. doi:10.1007/s10295-014-1576-3
- Singh, R. S., Kaur, N., and Kennedy, J. F. (2015). Pullulan and pullulan derivatives as promising biomolecules for drug and gene targeting. *Carbohydr. Polym.* 123, 190–207. doi:10.1016/j.carbpol.2015.01.032
- Singh, R. S., Saini, G. K., and Kennedy, J. F. (2008). Pullulan: microbial sources, production and applications. *Carbohydr. Polym.* 73, 515–531. doi:10.1016/j.carbpol.2008.01.003
- Sogutcu, E., Emrence, Z., Arikian, M., Cakiris, A., Abaci, N., Öner, E. T., et al. (2012). Draft genome sequence of *Halomonas smyrnensis* AAD6T. *J. Bacteriol.* 194(20), 5690–5691.
- Sohn, S. B., Kim, T. Y., Park, J. M., and Lee, S. Y. (2010). *In silico* genome-scale metabolic analysis of *Pseudomonas putida* KT2440 for polyhydroxyalkanoate synthesis, degradation of aromatics and anaerobic survival. *Biotechnol. J.* 5, 739–750. doi:10.1002/biot.201000124
- Staudt, A. K., Wolfe, L. G., and Shrout, J. D. (2012). Variations in exopolysaccharide production by *Rhizobium tropici*. *Arch. Microbiol.* 194, 197–206. doi:10.1007/s00203-011-0742-5
- Tang, X. Z., Kumar, P., Alavi, S., and Sandeep, K. P. (2012). Recent advances in biopolymers and biopolymer-based nanocomposites for food packaging materials. *Crit. Rev. Food Sci. Nutr.* 52, 426–442. doi:10.1080/10408398.2010.500508
- Teusink, B., van Enckevort, F. H., Francke, C., Wiersma, A., Wegkamp, A., Smid, E. J., et al. (2005). *In silico* reconstruction of the metabolic pathways of *Lactobacillus plantarum*: comparing predictions of nutrient requirements with those from growth experiments. *Appl. Environ. Microbiol.* 71, 7253–7262. doi:10.1128/JB.187.21.7254-7266.2005
- Teusink, B., Wiersma, A., Jacobs, L., Notebaart, R. A., and Smid, E. J. (2009). Understanding the adaptive growth strategy of *Lactobacillus plantarum* by *in silico* optimisation. *PLoS Comput. Biol.* 5:e1000410. doi:10.1371/journal.pcbi.1000410
- Thieme, F., Koebnik, R., Bekel, T., Berger, C., Boch, J., Büttner, D., et al. (2005). Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. *Vesicatoria* revealed by the complete genome sequence. *J. Bacteriol.* 187, 7254–7266. doi:10.1128/JB.187.21.7254-7266.2005
- Toksoy Oner, E. (2013). “Microbial production of extracellular polysaccharides from biomass,” in *Pretreatment Techniques for Biofuels and Biorefineries*, ed. Fang Z. (Berlin: Springer), 35–56.
- Valdés, J., Pedroso, I., Quatrini, R., Dodson, R. J., Tettelin, H., Blake, R., et al. (2008). *Acidithiobacillus ferrooxidans* metabolism: from genome sequence to industrial applications. *BMC Genomics* 9:597. doi:10.1186/1471-2164-9-597
- Van Tieghem, P. (1878). On sugar mill gum. *Ann. Sci. Nat. Bot. Biol. Veg.* 7, 180–203.
- Vorhölter, F. J., Schneiker, S., Goesmann, A., Krause, L., Bekel, T., Kaiser, O., et al. (2008). The genome of *Xanthomonas campestris* pv. *Campestris* B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. *J. Biotechnol.* 134, 33–45. doi:10.1016/j.jbiotec.2007.12.013
- Vroman, I., and Tighzert, L. (2009). Biodegradable polymers. *Materials* 2, 307–344. doi:10.3390/ma2020307
- Vu, B., Chen, M., Crawford, R. J., and Ivanova, E. P. (2009). Bacterial extracellular polysaccharides involved in biofilm formation. *Molecules* 14, 2535–2554. doi:10.3390/molecules14072535
- Wang, D., Chen, F., Wei, G., Jiang, M., and Dong, M. (2015). The mechanism of improved pullulan production by nitrogen limitation in batch culture of *Aureobasidium pullulans*. *Carbohydr. Polym.* 127, 325–331. doi:10.1016/j.carbpol.2015.03.079
- Wang, Q., Chen, X., Yang, Y., and Zhao, X. (2006). Genome-scale *in silico* aided metabolic analysis and flux comparisons of *Escherichia coli* to improve succinate production. *Appl. Microbiol. Biotechnol.* 73, 887–894. doi:10.1007/s00253-006-0535-y
- Whitney, J. C., and Howell, P. L. (2013). Synthase-dependent exopolysaccharide secretion in Gram-negative bacteria. *Trends Microbiol.* 21, 63–72. doi:10.1016/j.tim.2012.10.001

- Wibberg, D., Alkhateeb, R. S., Winkler, A., Albersmeier, A., Schatschneider, S., Albaum, S., et al. (2015). Draft genome of the xanthan producer *Xanthomonas campestris* NRRL B-1459 (ATCC 13951). *J. Biotechnol.* 204, 45–46. doi:10.1016/j.jbiotec.2015.03.026
- Williams, P. A., Phillips, G. O., and Seisun, D. (2007). *Overview of the Hydrocolloid Market*. Cambridge UK: RSC Publishing.
- Xu, C., Liu, L., Zhang, Z., Jin, D., Qiu, J., and Chen, M. (2013). Genome-scale metabolic model in guiding metabolic engineering of microbial improvement. *Appl. Microbiol. Biotechnol.* 97, 519–539. doi:10.1007/s00253-012-4543-9
- Yasar Yildiz, S., Anzelmo, G., Ozer, T., Radchenkova, N., Genc, S., Di Donato, P., et al. (2014). *Brevibacillus thermoruber*: a promising microbial cell factory for exopolysaccharide production. *J. Appl. Microbiol.* 116, 314–324. doi:10.1111/jam.12362
- Yasar Yildiz, S. Y., Radchenkova, N., Arga, K. Y., Kambourova, M., and Oner, E. T. (2015). Genomic analysis of *Brevibacillus thermoruber* 423 reveals its biotechnological and industrial potential. *Appl. Microbiol. Biotechnol.* 99, 2277–2289. doi:10.1007/s00253-015-6388-5
- Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A., Boldt, J., et al. (2011). Metabolic engineering of *Escherichia coli* for direct production of 1, 4-butanediol. *Nat. Chem. Biol.* 7, 445–452. doi:10.1038/nchembio.580
- Zakharevich, N. V., Averina, O. V., Klimina, K. M., Kudryavtseva, A. V., Kasianov, A. S., Makeev, V. J., et al. (2015). Complete genome sequence of *Bifidobacterium longum* GT15: identification and characterization of unique and global regulatory genes. *Microb. Ecol.* 70, 819–834. doi:10.1007/s00248-015-0603-x

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Ates. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Genome-Wide Transcriptional Response of *Saccharomyces cerevisiae* to Stress-Induced Perturbations

Hilal Taymaz-Nikerel, Ayca Cankorur-Cetinkaya and Betul Kirdar\*

Department of Chemical Engineering, Bogazici University, Istanbul, Turkey

## OPEN ACCESS

### Edited by:

Josselin Noirel,  
Conservatoire National des Arts et  
Métiers, France

### Reviewed by:

Trong K. Pham,  
University of Sheffield, UK  
Nuno Pereira Mira,  
Instituto Superior Técnico; Institute  
for Biotechnology and  
Bioengineering, Portugal

### \*Correspondence:

Betul Kirdar  
kirdar@boun.edu.tr

### Specialty section:

This article was submitted to  
Systems Biology,  
a section of the journal  
Frontiers in Bioengineering and  
Biotechnology

**Received:** 02 September 2015

**Accepted:** 04 February 2016

**Published:** 18 February 2016

### Citation:

Taymaz-Nikerel H, Cankorur-  
Cetinkaya A and Kirdar B (2016)  
Genome-Wide Transcriptional  
Response of *Saccharomyces*  
*cerevisiae* to Stress-Induced  
Perturbations.  
Front. Bioeng. Biotechnol. 4:17.  
doi: 10.3389/fbioe.2016.00017

Cells respond to environmental and/or genetic perturbations in order to survive and proliferate. Characterization of the changes after various stimuli at different -omics levels is crucial to comprehend the adaptation of cells to the changing conditions. Genome-wide quantification and analysis of transcript levels, the genes affected by perturbations, extends our understanding of cellular metabolism by pointing out the mechanisms that play role in sensing the stress caused by those perturbations and related signaling pathways, and in this way guides us to achieve endeavors, such as rational engineering of cells or interpretation of disease mechanisms. *Saccharomyces cerevisiae* as a model system has been studied in response to different perturbations and corresponding transcriptional profiles were followed either statically or/and dynamically, short and long term. This review focuses on response of yeast cells to diverse stress inducing perturbations, including nutritional changes, ionic stress, salt stress, oxidative stress, osmotic shock, and to genetic interventions such as deletion and overexpression of genes. It is aimed to conclude on common regulatory phenomena that allow yeast to organize its transcriptional response after any perturbation under different external conditions.

**Keywords:** yeast, transcriptome, perturbation, regulation, stress

## INTRODUCTION

The central challenge in systems biology is to construct the whole life model for the prediction of cellular response to the changing environments. Therefore, the genome-level understanding of the cellular response to both genetic and environmental perturbations is extremely important in modeling. Consequently, systematic perturbation experiments were conducted to reach this goal. *Saccharomyces cerevisiae* has been investigated in biochemical and genetics laboratories for many decades, since it is considered to be a good model organism. Cellular response of this organism at different -omics levels to different perturbations has also been extensively studied by systematically introducing environmental changes. The availability of a deletion collection made it also an attractive organism to screen the cellular response to deletions of specific genes. The systems biology-based studies resulted in the understanding of several pathways and the cellular behavior of the yeast cells to different perturbations. Although it is probably the best-understood organism, we are still far away from modeling the response of this organism to perturbations (Boone, 2014).

Eukaryotic cells reprogram the expression of those genes that are essential for adapting to the changing conditions as a response to perturbation. Microarray technology allowed investigation

of expression profiles of thousands (whole-genome arrays) or hundreds (low-density arrays) of genes simultaneously. ArrayExpress at the European Bioinformatics Institute (EBI) and the Gene Expression Omnibus (GEO) database at the National Center for Biotechnology Information (NCBI) are the two major public databases of microarray data (Barrett et al., 2013; Kolesnikov et al., 2015). Although they have different designs, both databases support the minimum information about a microarray experiment (MIAME), a standard guideline for describing a microarray experiment (Brazma et al., 2001). The establishment of RNA-seq, providing detailed measurement, and lower technical discrepancy, became another attractive analytical tool in transcriptomics (Nagalakshmi et al., 2008; van Dijk et al., 2011; Nookaew et al., 2012). Most of the high-throughput data have been generated by sampling the experimental system at single-time point. Generation of the time-series gene expression data is considered to be very important to understand the dynamic nature of biological systems. The efforts to study and model such dynamic data were reviewed in detail by Bar-Joseph et al. (2012), Yosef and Regev (2011), and Secrier and Schneider (2014).

The programming of gene expression in cells occurs on a broad range of time-scales from rapid responses (minutes to hours of response to environmental stresses) to slower (hours to days during development) processes (Yosef and Regev, 2011). An early analysis of the transcriptomic response of yeast cells to diverse environmental changes indicated that a large set of genes (~900) showed a comparable and severe response to different perturbations sensing them as an environmental stress [environmental stress response (ESR)]. The upregulated genes were related to stress defense regulated by Msn2p and Msn4p and downregulated genes were associated with ribosome biogenesis and protein synthesis. An important observation to be noted was the involvement and the regulation of different isoenzymes in response to several perturbations (Gasch et al., 2000; Causton et al., 2001; Gasch, 2003). Positive and negative regulators of protein kinase A (PKA) pathway were also reported to be induced within the ESR genes (Gasch et al., 2000). This common response to perturbations was reported to be important for preparing cells in response to further possible changes in the environment (Berry and Gasch, 2008; Mitchell et al., 2009) leading to stress (cross-) tolerance. However, it should be noted that the correlation between upregulation of gene expression and its requirement for fitness was not always significant (Giaever et al., 2002; Giaever and Nislow, 2014).

Zakrzewska et al. (2011) used yeast haploid deletion collection to find the genes and their related functions involved in the stress survival and in the gain of stress tolerance. Survival analysis of the yeast cells revealed that a general stress response (GSR) increases survival after a period of mild stress pretreatment, and this survival of stress was negatively correlated with mutant growth rate. Resistance to stress and corresponding tolerance gained, induced by severe perturbations in *S. cerevisiae*, are directed by specific processes for each stress and at the same time by general processes for all stresses. Growth rate was acknowledged for being responsible for tolerating stress and growth rate reduction was found to be responsible for gaining tolerance to stress. Transcriptomic analysis in response to several perturbations in chemostats at different growth rates indicated that there was a notable correlation

between the growth rate-dependent genes and ESR genes (Gasch et al., 2000; Regenberget al., 2006; Castrillo et al., 2007; Brauer et al., 2008; Fazio et al., 2008). It has been proposed that most of the alterations in the expression of ESR genes are associated with the lowered growth rate as a result of a change in the environment. Msn2p and Msn4p are found to be important in the organization of transcriptome, but not in the observed changes, in the phenotype due to changes in the growth rate (Zakrzewska et al., 2011). Components of stress-specific signaling pathways and effectors of these pathways intervening functional adaptation in response to changes in environmental conditions have been identified (Chasman et al., 2014). An extensive review of signaling pathways that control proliferation versus stress defense in yeast and mammalian cells was recently provided (Ho and Gasch, 2015).

Transcriptomic data obtained from perturbation experiments are an extremely important step in the construction of large scale, system-based models that can be used to predict the cellular response to perturbations. The integration of this data with other -omics, pathway information, computational, and statistical tools, is crucial to improve the accuracy in the prediction capabilities of the models. However, it is important to note that these experiments should be carefully designed and interpreted according to the needs of the investigators. In the present review, systematically introduced perturbations to monitor the whole-genome level transcriptomic response of *S. cerevisiae* cells to these perturbations will be summarized. Most commonly studied stress-causing perturbations, such as the changes in the types and quantities of available nutrients, oxidative reagents, temperature, osmolarity, and the metal ions, were selected to review. Moreover, these perturbations are closely related to the optimization of industrial applications and to human diseases. Although we review each perturbation in a different section, it is not always possible to analyze only one stress factor alone. For example, during fermentation, *S. cerevisiae* wine strains undergo considerable stress due to the high concentrations of sugars, producing high osmotic pressure, followed by ethanol accumulation, addition to nitrogen limitation, low pH, and the presence of SO<sub>2</sub>, all imposing pressure on the cells (Treu et al., 2014).

## TRANSCRIPTOMIC RESPONSE TO NUTRITIONAL CHANGES

The ability of living organisms to acclimate to alterations in their nutritional environment is necessary for their survival, and they have developed mechanisms to cope with the new conditions quickly and effectively. Yeast cells sense the nutrients in the environment *via* a group of major nutrient-signaling pathways, which coordinate general responses, such as cellular proliferation and stress resistance. Systematic perturbation experiments were performed to understand the underlying molecular mechanisms responsible for this adaptation process to the changes in the environment. The response and tolerance to nutrient-relevant stresses have been largely reviewed by Teixeira et al. (2011) and Conrad et al. (2014). Our focus here will be to review the selected studies on the transcriptomic response of *S. cerevisiae* to nutrient limitations and to transient changes in the nutrient environment

as well as on the studies aiming to elucidate the nutrient-signaling pathways.

## Response to Nutrient Limitations

Nutrient limitation has drawn the attention of researchers, because microorganisms encounter nutrient limitation in their natural environments and during industrial processes. The first genome-wide transcriptome analysis in *S. cerevisiae* by DeRisi et al. (1997) investigated the response of yeast cells at transcriptomic level during the diauxic shift using microarrays. This pioneering work indicated that the passage from a glucose rich medium to a glucose-depleted medium involves the integration of a number of major signaling and regulatory pathways. These environmental perturbation experiments revealed that glucose depletion leads to the induction of cytochrome *c*-related genes and the genes involved in the TCA/glyoxylate cycle and carbohydrate storage and to the repression of those involved in protein synthesis, including ribosomal proteins, tRNA synthetases, and translation, elongation, and initiation factors. They further investigated the genome-wide transcriptional response to the deletion of *TUP1*, general transcription repressor, and to the overexpression of *YAP1*, transcription activator, to understand the contribution of those individual regulatory proteins to the reprogramming of transcriptional response to glucose. These genetic perturbation experiments revealed the group of genes, expression of which was altered by the deletion of these transcription factors (DeRisi et al., 1997). These results demonstrated for the first time that microarray technology is a useful tool for the genome-wide exploration of expression patterns of genes upon environmental and genetic perturbations. Transcriptional alterations resulting from deletion or overexpression of regulatory molecules identified by this approach were used to dissect and characterize regulatory pathways and networks.

Gasch et al. (2000) examined the temporal transcriptomic changes in yeast cells in batch cultures exposed to amino acid starvation or to nitrogen depletion for a period of 6 h and reported the alteration in the expression levels of ESR genes under both conditions. The genes involved in carbohydrate metabolism, detoxification of reactive oxygen species (ROS), cellular redox reactions, cell wall modification, protein folding and degradation, DNA damage repair, fatty acid metabolism, metabolite transport, vacuolar and mitochondrial functions, autophagy, and intracellular signaling were found to be induced and the genes related to growth-related processes and ribosome biogenesis were reported to be repressed. Furthermore, the transcriptomic analysis of the strains carrying deletion of *Msn2p*, *Msn4p*, and both of these factors revealed that the majority of ESR genes were under the control of *Msn2p* and *Msn4p*. Starvation-specific response of yeast cells was reported to be related to switch from active growth to a growth-arrested state.

The observation of the effect of changing growth rates on the transcriptomic response to changes in the nutrient environment resulted in a shift of mode of fermentations from batch to chemostat with a constant growth rate. Boer et al. (2003) carried out the analysis of the transcriptome of *S. cerevisiae* cells under carbon-, nitrogen-, phosphorus-, or sulfur-limitation, at a dilution rate of  $0.1\text{ h}^{-1}$ . This study revealed the significant alterations in the

expression levels of 31% of genes between at least two growth conditions. The genes involved in the uptake and phosphorylation of glucose, uptake and metabolism of fatty acids and storage carbohydrates, glyoxylate cycle, and gluconeogenesis, uptake and utilization of alternative carbon sources were upregulated under carbon limitation in addition to the induction of the few genes involved in the protection against oxidative stress. The expression levels of the genes involved in the transport such as low or moderate affinity hexose transporters, the genes involved in the glucose repression, cell proliferation, and differentiation were downregulated under glucose limitation when compared with three other conditions. Under nitrogen limitation, the genes involved in the metabolism of nitrogen-containing compound were observed to be upregulated, and the promoter analysis of coregulated genes revealed that *Gln3p*, *Gat1p*, *Dal80p*, and *Gzf3p* have important roles in the regulation of genes under nitrogen catabolite repression. Under phosphate limitation, the genes implicated in the uptake and metabolism of inositol phosphate, in the phosphate metabolism, and in the process of phosphorylation of metabolites were found to be induced. From promoter analysis, *Pho4p* was predicted to be the regulator of these events. Under sulfur limitation, the upregulation of the genes associated with the uptake of sulfate and sulfur-containing molecules and with the sulfur assimilation process was observed. Promoter analysis predicted that the transcription factors including *Chf1p*–*Met4p*–*Met28p* complex, *Met31p*, and *Met32p* are involved in the regulation of these genes. The genes involved in the glycogen metabolism, in the export of sulfite and in copper transportation were found to be repressed under phosphate limitation, and the majority of these downregulated genes were predicted to be dependent on *Msn2p* and *Msn4p*. It has been suggested that the cellular metabolism reorganized to encounter the needs for nutrient-limited growth of the yeast cells (Boer et al., 2003). Wu et al. (2004) have also examined the genome-wide transcriptome of yeast under five (glucose, ethanol, ammonium, phosphate, and sulfate) different nutrient-limited conditions at a dilution rate of  $0.1\text{ h}^{-1}$ . The genes affected under each condition were identified with the comparison to its corresponding steady state. This study revealed that the genes involved in the TCA cycle, oxidative phosphorylation, and the genes encoding high-affinity glucose transporters were upregulated under glucose-limited condition. The genes involved in glyoxylate cycles, gluconeogenesis, and nitrogen catabolite repression were upregulated under ethanol-limited conditions even in the presence of sufficient levels of nitrogen in the medium. An interesting result from this study was the observation of the activation of iron-associated genes, including *FTH1*, *FET3*, *FRE3*, *FIT2*, and *FIT3*, by glycolysis. The variances between these two investigations may stem from the differences in the strain background, in the experimental conditions, and in the analysis of data.

The relationship between growth rate and response to nutrient limitations draws the attention of many researchers (Castrillo et al., 2007; Boer et al., 2008, 2010; Brauer et al., 2008; Slavov and Botstein, 2011, 2013). Castrillo et al. (2007) have investigated the effect of the growth rate on the transcriptional, metabolic, and proteomic responses to nutritional changes in yeast with a special emphasis on TOR pathway. The experiments were carried out at

growth rates of 0.07, 0.1, and 0.2 h<sup>-1</sup> under four different nutrient limitations (glucose, ammonium, phosphate, and sulfate) in chemostats. Data analysis was performed by comparing the nutrient limited with nutrient sufficient condition in each case. This study revealed a common response to increase the growth rates in yeast at all omics levels under all nutrient limiting conditions. The genes upregulated with the increasing growth rate and independent of the nutritional limitation were found to be involved in the processes related to translation initiation, ribosome biogenesis, protein biosynthesis, RNA metabolism, nucleic acid metabolism, nucleus import/export, and proteasome function. The downregulated genes were implicated in biological processes related to the response to external stimulus, cell transduction, autophagy, homeostasis, response to stress, and vesicle recycling within Golgi. It has been suggested that these groups of genes might be involved in maximizing the efficient use of cellular resources in each limiting condition at each different growth rate. Furthermore, it has been demonstrated in this study that most of the genes regulated with growth rates are also targets of the TOR signaling pathway. The models to explain the regulation of metabolic fluxes of carbon/nitrogen, sulfur/folate, and leucine biosynthesis at the protein level were constructed through the integration of quantitative proteomic and metabolomic data.

Brauer et al. (2008) also studied the relationship between the growth rate, entry into cell cycle, glucose metabolism, and the transcriptomic response to different nutritional limitations (i.e., glucose, ammonium, sulfate, phosphate, uracil, or leucine) in 36 steady-state conditions. The transcript levels of ~27% of all yeast genes were found to be either positively or negatively correlated with growth rate, independent of any nutrient limitation, and provided further support for the results published by Castrillo et al. (2007). The genes involved in ribosomal functions and stress response were found to be positively correlated with growth rate, whereas the genes involved in peroxisomal functions were identified to be negatively correlated with growth rate. The relationship between growth rate and cell-cycle population in G0/G1 phase was also linear independent of limiting nutrient. They have also constructed a linear model, based on the gene expression values, to predict relative “instantaneous growth rate” in both batch and continuous cultures. The observation of the complete consumption of glucose and failure in arresting in cell cycle, when the growth of the auxotrophic strains is limited by the auxotrophic requirements, led researchers to investigate this phenomenon further.

Analysis of the transcriptomic response of auxotrophic yeast cells grown in chemostat cultures limited only with their respective auxotrophic requirements, at different growth rates, indicated a decoupling of growth rate response (GRR) from nutrient signaling, and the magnitude of the response was found to be less than that recorded in prototrophs. It has been suggested that growth rate signal is important in the regulation of fermentation/respiration, the GRR, and the cell division cycle (Brauer et al., 2008; Slavov and Botstein, 2013).

## Response to Transient Nutrient Changes

The reorganization of the transcriptomic and/or metabolomic response to transient changes in the nutrient environment was

investigated by the impulse-like addition of the limited nutrient into yeast cells grown in chemostats under a specific nutrient limitation. The dynamic adaptation to the changing concentration of the nutrient was followed in time-course experiments after the perturbation of the steady state by the addition of the limited nutrient.

Kresnowati et al. (2006) studied transcriptional and metabolic response of yeast cells within the initial 5 min of the perturbation when aerobic, tightly controlled glucose-limited chemostat cultures were subject to a glucose pulse, and a new insight could be provided for the temporal organizations of metabolic and transcriptional events. This study provided the first example of the integration of transcriptome with metabolic changes during the initial response (minutes) to a well-defined perturbation of *S. cerevisiae* cells. They have shown that although metabolic and transcriptomic data are well correlated, the changes in metabolite concentrations were observed in seconds, and the changes in the transcriptomic response were detected later. Cells have responded to this increasing glucose level at transcriptional and metabolic level in order to adapt to the change from respiratory to respiro-fermentative metabolism and to growth acceleration. The same group has also studied the long-term response of yeast cells to a glucose pulse and depletion of oxygen by investigating the dynamic adaptation of *S. cerevisiae* grown in aerobic, glucose-limited chemostat cultures to an anaerobic pulse of glucose for a longer period of time (5, 10, 30, 60, and 120 min). The fact that one-third of the genes were significantly and differentially expressed indicated an important reprogramming and strong response of the cells after the pulse. Most of these genes were found to be related to growth and carbon catabolite repression. The expression levels of several genes, including the genes encoding ribosomal biogenesis and ribosomal proteins, were changed in opposite direction 30 min after the pulse (van den Brink et al., 2008).

Bradley et al. (2009) have also shown the coordinated changes between the transcriptome and metabolome by analyzing the dynamic transcriptomic and metabolomic response of yeast cells grown on filters to glucose and ammonium starvation and developed an integrative approach based on Bayesian framework to predict metabolite–transcript correlations.

Dikicioglu et al. (2011, 2012) investigated the transcriptional response in the short (seconds) and long term (hours), by introducing a glucose pulse to a glucose-limited culture and an ammonium pulse to an ammonium-limited culture of yeast cells. The integration of the transcriptomic data from these perturbation experiments with two different network-based approaches presented further information about the time-based regulation of transcriptional reorganization of yeast cells under these conditions. Integration of this transient transcriptome data with the corresponding metabolome data combined with metabolic pathway information revealed the whole-genome level long-term reorganization of the yeast cells. The changes in the transcriptome and metabolome from an initial limited state, followed by a sudden pulse and then by returning to the limited state were identified for the first time in this study. The transient organization of the *de novo* biosynthetic pathways and salvage pathways under these conditions was reported through integration of transcriptome and metabolome.



## Nutrient Signaling

The sensing and signaling networks that direct the transcriptional and metabolic reorganization in response to the changing nutritional conditions were subject to several studies and reviewed largely by Zaman et al. (2008). In order to understand the regulation of transcriptional response of yeast cells to glucose, the transcriptomic data of the selected yeast strains that carry conditional alleles in response to glucose perturbation were analyzed. The integration of genetic knowledge with microarray analysis revealed the presence of five interlocking pathways in glucose signaling. PKA and PKB (Sch9p) were found to be the main regulators, whereas Snf1 and Rgt pathways have restricted roles in the regulation of the transcriptomic response to glucose. A schematic model illustrating the glucose signaling network was constructed. The perception of the environmental nutrients by the cell was reported to be the main determinant of growth rate-dependent organization of transcriptomic response (Zaman et al., 2009).

Livas et al. (2011) investigated the genome-wide transcriptional response of wild-type yeast cells and two suppressor strains lacking PKA activity when grown on ethanol and exposed to glucose for 30 min. This study revealed the set of genes, which are induced or repressed by glucose either dependent or independent of PKA activity. Induction of the genes involved in glucose metabolism and repression of the genes involved in the utilization of alternative carbon sources were found to be exclusively controlled by other pathways and independent of PKA. The genes involved in amino acid synthesis were found to be regulated by PKA, and the response to glucose is transient. The genes regulated by redundant PKA-dependent and PKA-independent signaling pathways included the genes encoding glucose transporters, ribosomal proteins, and the genes involved in the utilization of non-fermentable carbon sources. A fourth group of genes was found to be regulated by the cooperative signaling pathways, including PKA. This study provided additional support that PKA signaling is an important player in the reorganization of the transcriptomic response of *S. cerevisiae* to glucose.

Hughes Hallett et al. (2014) examined the TOR kinase complex I (TORC1) pathway, which has a key role in the regulation of cell growth, by investigating the changes in the yeast transcriptome in response to various perturbations, including glucose starvation and nitrogen starvation for 20 min. Under glucose starvation, PPA2 branch of TORC1 pathway has a low level activity, whereas Sch9 branch was completely switched off. The transcriptomic analysis of *SNF1* deleted strains indicated that cAMP activated kinase (Snf1p) represses TORC1–Sch9 signaling but hyper-inactivates TOR–PPA2 signaling indicating that Snf1p regulates the TORC1 pathway. However, the presence of additional factors remains to be elucidated. Under nitrogen starvation, TORC1–Sch9 signaling was found to be inactive, whereas TORC1–PPA2 signaling has a high activity. The analysis of the corresponding deletion strains showed that the regulators of the TORC1, Gtr1p/Gtr2p, Npr2p/Npr3p, and Rho1p have important roles in nitrogen or amino acid starvation.

In a recent study, Oliveira et al. (2015) investigated the dynamic regulation of nitrogen metabolism by TORC1 pathway in yeast cells by analyzing transcriptome, proteome, and metabolome data. Codesigning of a perturbation matrix to follow the changes

at different omics levels and the use of probabilistic model-based analysis that incorporates the prior knowledge resulted in the identification of putative targets and inputs of TORC1 including a novel putative glutamine signal.

In order to understand the pleiotropic role of the PAS kinase Rim15p in the integration of nutrient-signaling networks, such as TOR, PKA, and the Pho80–Pho85 kinase pathways, the transcriptomic responses of prototrophic *RIM15* deletion mutant and a congenic *RIM15* reference strain of *S. cerevisiae* were comparatively analyzed. Cells were grown under severe calorie restriction in anaerobic retentostat cultures near non-growing conditions. This study revealed the important function of Rim15p in cell-cycle arrest and in the integration of nutrient sensing and signaling pathways under nutrient-depleted conditions (Bisschops et al., 2014).

The details of all these nutritional perturbation experiments in *S. cerevisiae* cells are summarized in **Table 1**. Many aspects of the response to limitation of C-, N-, S-, and P-sources could be revealed, and the details of the growth-regulated response could be elucidated as well as a wealth of information about transcriptional reorganization in response to the changing nutritional environment that was provided by the above mentioned studies. It is important to keep in mind that the information coming from different studies is far away from being integrated into a model due to the differences in the strain background or in the experimental design. Moreover, the use of different algorithms to analyze and interpret the data provokes an additional difficulty in the comparison or integration of these findings. The challenge continues to identify the nutrition signaling pathways by specifically designed experiments on the basis of these observations. It is important to design an experimental platform to analyze and integrate the response to nutritional changes at different omics levels from wild-type as well as from specifically selected strains, based on the previous knowledge.

## TRANSCRIPTOMIC RESPONSE TO OXIDATIVE PERTURBATIONS

Exposure to oxidative agents causes the production of ROS, which are known to cause severe cellular damage and be involved in cellular processes, such as aging and apoptosis, as well as in the molecular pathogenesis of several severe disorders. The detailed information on cellular antioxidant systems, which protect yeast cells against ROS accumulation, can be found in the reviews (Farrugia and Balzan, 2012; Morano et al., 2012). Oxidative agents used to investigate oxidative stress response (OSR) are hydrogen peroxide, lipid hydroperoxides, organic hydroperoxides, such as cumene hydroperoxide (CHP), linoleic acid hydroperoxide, superoxide anion, and heavy metals, such as Fe<sup>2+</sup> and Cd. A systematic analysis of yeast deletion mutants revealed that the genes involved in the OSR, induced by different agents, consists of a set of genes known as “core genes,” which are observed under a wide variety of oxidative stress condition and another group of genes known as the “oxidant specific” genes (Thorpe et al., 2004). Several perturbation studies were carried out to investigate the dynamic transcriptomic response of yeast cells to oxidative stress

**TABLE 1 | Nutritional perturbation experiments in *S. cerevisiae* cultures.**

Perturbation	Experiment type	Exposure (sampling) time	Measurement technique	Reference
<b>Nutrient limitations</b>				
From glucose rich to glucose-depleted medium (Diauxic shift)	NA	From 9 to 21 h	cDNA microarrays	DeRisi et al. (1997)
From glucose grown to amino acid starvation	Shake flasks	Sampled at early exp. phase (0), 0.5, 1, 2, 4, and 6 h	cDNA microarrays	Gasch et al. (2000)
From glucose grown to amino acid starvation + nitrogen limitation		Sampled at early exp. phase (0), 0.5, 1, 2, 4, 8, 12 h, 1, 2, 3, and 5 days		
Aerobic carbon, nitrogen, phosphorus, or sulfur limited	Chemostat at 0.1 h <sup>-1</sup>	Sampled at steady state	Affymetrix Yeast Genome S98 Array	Boer et al. (2003)
Glucose, ethanol, ammonium, phosphate, or sulfate limited	Chemostat at 0.1 h <sup>-1</sup>	Sampled at steady state and starvation phase	Arrays supplied by Hoheisel	Wu et al. (2004)
Carbon, nitrogen, phosphorus, or sulfur limited	Chemostat at 0.07, 0.1, and 0.2 h <sup>-1</sup>	Sampled at steady state	Affymetrix Yeast Genome S98 Array	Castrillo et al. (2007)
Glucose, ammonium, sulfate, phosphate, uracil, or leucine limited	Chemostat at 0.05, 0.1, 0.15, 0.2, 0.25, and 0.3 h <sup>-1</sup>	Sampled at steady state	Agilent Yeast Oligo Microarray (V2)	Brauer et al. (2008)
Histidine-, lysine-limited	Chemostat at 0.05, 0.1, 0.15, 0.2, 0.25, and 0.3 h <sup>-1</sup>	Sampled at steady state	Agilent Yeast Oligo Microarray	Slavov and Botstein (2013)
<b>Transient nutrient changes</b>				
Glucose (1 g/L) pulse to glucose-limited	Chemostat at 0.05 h <sup>-1</sup>	0, 0.5, 1, 2, 3.5, 5, and 5.5 min	Affymetrix Yeast Genome S98 Array	Kresnowati et al. (2006)
Anaerobic glucose (200 mM) pulse to aerobic glucose-limited	Chemostat at 0.1 h <sup>-1</sup>	0, 5, 10, 30, 60, and 120 min	Affymetrix Yeast Genome S98 Array	van den Brink et al. (2008)
Switch to glucose-free or ammonium-free medium	Filter-cultures	Sampled at exp. phase (0), 10, 30, 60, 120, 240, and 480 min	Agilent Yeast Oligo Microarray (V2)	Bradley et al. (2009)
Glucose (2% w/v) pulse to glucose-limited and ammonium (0.3% w/v) pulse to ammonium-limited	Chemostat at 0.1 h <sup>-1</sup>	Sampled at 0, 20, 40, 60 s, 8, 16, 24, 32min, 1, 2, 3, 4, 5, and 7 h	Affymetrix Yeast Genome 2.0 Array	Dikicioglu et al. (2011, 2012)
<b>Nutrient signaling</b>				
Galactose (2%) or glucose (2%) addition to glycerol grown cells	NA	Sampled at OD <sub>600</sub> = 0.25 (0), 20, 40, 60, and 80 min	Agilent Yeast Oligo Microarray	Zaman et al. (2009)
Glucose (4%) addition to ethanol grown cultures	NA	Sampled at early exp. phase (0), and 30 min	Affymetrix Yeast Genome S98 Array	Livas et al. (2011)
Switch to glucose, nitrogen, or amino acid-free medium	Shake flasks	Sampled at OD <sub>600</sub> = 0.60 (0), 20 min	Agilent Yeast Oligo Microarray (V2)	Hughes Hallett et al. (2014)
Glutamine (400 mg/L) pulse to proline grown cultures from proline + glutamine grown cultures to glutamine depletion	Batch reactor	Sampled at 0, 3, 7, 10, 14, 24, 56, and 120 min	Affymetrix Yeast Genome 2.0 Array	Oliveira et al. (2015)
Anaerobic glucose limited	Chemostat at 0.025, 0.05, 0.1 h <sup>-1</sup>	Sampled at steady state	Affymetrix Yeast Genome S98 Array	Bisschops et al. (2014)
	Retentostat	Sampled at 2, 9, 16, and 20 days		

caused by different oxidative agents. Since the hydrogen peroxide is a natural ROS as a by-product of the aerobic metabolism and was most extensively used as a model system for oxidative stress, we will mainly focus in this review on the studies where the oxidative stress was induced by hydrogen peroxide treatment.

The dynamic transcriptomic response to the addition of hydrogen peroxide and menadione into batch cultures was first investigated by Gasch et al. (2000) among other environmental stresses for a period of 2–3 h. The genes encoding superoxide dismutases, glutathione peroxidases, thiol-specific antioxidants, thioredoxin, thioredoxin reductases, glutaredoxin, and glutaredoxin reductase in addition to ESR genes were found to be specifically upregulated in response to both chemicals. Transcriptomic response of *YAP1* deleted yeast strain indicated that Yap1p is an important regulator of the OSR. Causton et al. (2001) have also investigated the transcriptomic response of yeast cells to hydrogen peroxide similarly and reported the specific upregulation of *ROX1*, which is a

repressor of hypoxic genes. Comparative transcriptomic analysis of the wild-type, *yap1Δ*, *yap2Δ*, *yap1Δ yap2Δ*, yeast cells after treatment with hydrogen peroxide showed that these proteins are regulators of different biological processes in OSR (Cohen et al., 2002).

TOR kinase complex I pathway is known to be involved in the response to a vast variety of stresses (Loewith and Hall, 2011). Hughes Hallett et al. (2014) examined the TORC1 pathway process information by investigating transcriptional reorganization of yeast cells in response to various perturbations, including oxidative stress induced with hydrogen peroxide using an integrative approach. The analysis of two different modules indicated that the genes in the TORC1–Sch9 pathway were downregulated and the expression levels of the genes in the TORC1–PP2A were not or little affected under oxidative shock. The investigation of the phosphorylation levels of the proteins regulated by PP2A by bandshift assay showed that Npr1p and Gln3p are dephosphorylated but

Nnk1p remained phosphorylated upon oxidative stress. These results indicated that TORC1–PP2A-branch signaling is weak or moderate under oxidative stress. Furthermore, the analysis of three yeast strains that stop the transmission of the signal from Npr2p/Npr3p, Gtr1p/Gtr2p, and Rho1p to TORC1–Sch9 signaling and *SNF1* deleted strains resulted in the observation that these proteins do not have any effect on this signaling under oxidative stress as well as under osmotic and heat stresses.

Investigation of the dynamic transcriptional response of yeast cells to oxidative stress induced by the addition of CHP at mid-exponential phase under fully controlled conditions revealed early transcriptional events (Sha et al., 2013). Approximately 54% of the genes that are regulated by Msn2p/Msn4p were also found to be significantly and differentially expressed after the treatment with CHP. *YAP1* was found to be upregulated after 6 and 20 min of induction and 52% of its targets were differentially and significantly expressed upon oxidative stress. *HMS2*, *MET28*, *YAP5*, *NUT2*, *ROX1*, and *SUT2* encoding transcriptional factors were also found to be induced during the early response within the first 6 min after the addition of CHP. *MET1*, *MET12*, *MET16*, *MET22*, *MET3*, *MET8*, *CYS3*, and *STR3* regulating sulfur metabolism, which are targets of *MET28*, were upregulated within 20 min. Other members of the YAP family (*YAP3*, *YAP5*, and *YAP7*) were also upregulated in the early response to CHP. Drug resistance-related proteins, the proteins involved in cell wall and cytoskeleton metabolism, and another group of genes of unknown function were found significantly induced within 6 min after the stress induction and returned immediately to their basal levels. A transient repression of the genes involved in cell growth, DNA replication, transcription, and translation was observed within this interval. The genes associated with mitochondrial function and vesicle trafficking were also transiently downregulated in this early period. The transcript levels of the genes that are involved in glutathione, glutaredoxin, and thioredoxin systems and the genes encoding ROS removing enzymes were induced within this early period. The genes encoding the enzymes of the oxidative branch of the pentose phosphate pathway were upregulated, whereas the branch leading to the synthesis of nucleic acids was repressed. The transcription factor Rpn4p that is regulator of the synthesis of the proteasome subunits was induced earlier than the genes encoding of these subunits to cope with the removal of the accumulated oxidized proteins. A comparative analysis of the results with previous studies (Gasch et al., 2000; Causton et al., 2001), whereas the hydrogen peroxide was used to induce the oxidative stress, revealed the induction of the ESR genes, and the genes involved in the glutathione metabolism and the pentose phosphate pathway. The biological processes associated with transcription and translation were downregulated in all three studies. The induction of cell wall and membrane was specific for the antioxidant CHP. The genes associated with mitochondrial processes were downregulated in response to CHP, whereas upregulated in response to hydrogen peroxide (Sha et al., 2013).

In a recent study, Zhao et al. (2015) investigated the transcriptional response of a strain, which has higher peroxide tolerance ability to 2 mM hydrogen peroxide exposure in comparison to that of control strain. The genes involved in carbohydrate metabolism, fatty acid degradation, glycolysis/gluconeogenesis, peroxisomal

matrix, pyruvate metabolism, amino acid metabolism, and nucleotide repair pathways were found to be significantly and differentially expressed between two strains in response to hydrogen peroxide. MAP kinase and cAMP–PKA signaling pathways, which were significantly enriched by the genes responsive to oxidative perturbation, were identified to be involved in the oxidative tolerance of the mutant strain.

A whole-genome scale analysis at different omics levels revealed that Slf1p, which is La-related protein, is involved in the translational control of oxidative stress induced by hydrogen peroxide (Kershaw et al., 2015). Deletion and mutated strains of *SLF1* were used in this study and cultures were treated with hydrogen peroxide for 10 or 60 min to induce oxidative stress.

The details of all these oxidative perturbation experiments in *S. cerevisiae* cells are summarized in Table 2. These studies, focused on the response of *S. cerevisiae* to oxidative stress, revealed a set of regulators including Yap1p and its homologs, Skn7p, Msn2p, and Msn4p and their selective targets. However, construction of a quantitative model, which incorporate sensing, signaling, and regulation, could not be constructed. It should also be noted that further studies are required to identify missing information coming from a unique experimental platform and carefully designed perturbation experiments.

## TRANSCRIPTOMIC RESPONSE TO THE CHANGING TEMPERATURES

Yeast cells encounter rapid and large differences in temperature in nature, and industrial strains are being optimized to grow at certain temperatures, depending on the production process. Therefore, in order to understand the mechanisms for adapting to and showing tolerance to different temperatures, perturbation experiments were designed and performed. The adaptation of wine strains to cold is also important to improve the aroma of the wine. Yeast strains were stored at freezing temperatures, and it is critical to comprehend storage temperature effects on the viability/physiology of the cells. Therefore, the transcriptional response of yeast cells to different temperatures was extensively studied.

### Heat Shock or Adaptation to Higher Temperatures

Temporal genome-wide transcriptional changes of wild-type and *MSN2/MSN4* deleted yeast cells in response to a temperature shift from 25 to 37°C was first investigated by Gasch et al. (2000). The changes in the expression levels of ESR genes were observed within the first minutes after the heat shock, and the majority of these genes were found to be regulated by Msn2p/Msn4p by the investigation of the transcriptomic response of double deleted mutants to heat shock. The genes involved in protein folding chaperons were observed to be induced later. Analysis of the transcriptomic response to a similar temperature shift has also provided further support for these findings (Causton et al., 2001).

A comparative genome-wide analysis including the transcriptional analysis of wild-type and *rpdl3Δ* strains grown at 25 and 39°C for 20 min resulted in the finding that RpdL3 histone deacetylase complex is involved in the partial regulation of gene

**TABLE 2 | Oxidative perturbation experiments in *S. cerevisiae*.**

Perturbation	Experiment type	Exposure (sampling) time	Measurement technique	Reference
Hydrogen peroxide (0.30 mM) addition at early exp. phase	Shake flasks	0, 10, 20, 30, 40, 50, 60, 80, 100, and 120 min	cDNA microarrays	Gasch et al. (2000)
Hydrogen peroxide (0.40 mM) addition at mid-exp. phase	Shake flasks	0, 10, 20, 40, 60, and 100 min	Affymetrix YE6100 gene chips	Causton et al. (2001)
Hydrogen peroxide (0.6 mM) addition at mid-exp. phase	NA	1 h	Oligonucleotide arrays	Cohen et al. (2002)
Hydrogen peroxide (2 mM) addition at mid-exp. phase	Shake flasks	0, 20 min	Agilent Yeast Oligo Microarray (V2)	Hughes Hallett et al. (2014)
CHP (190 $\mu$ M) addition at mid-exp. phase	Batch bioreactor	0, 3, 6, 12, and 20 min	Affymetrix Yeast Genome S98 Array	Sha et al. (2013)
Growth with 2 mM hydrogen peroxide	Shake flasks	Sampled at mid-exp. phase	RNA-sequencing	Zhao et al. (2015)
Hydrogen peroxide (0.4 mM) addition at mid-exp. phase	NA	0, 15, and 60 min	RNA-sequencing	Kershaw et al. (2015)

expression upon heat exposure. This study indicated the important role of the chromatin modifications in the reorganization of transcriptomic response to the heat exposure. Hsf1p and Msn2p/Msn4p are main regulators of the heat shock, and this complex has a role in the activation of the genes regulated by Msn2p/Msn4p (Ruiz-Roig et al., 2010).

Mensonides et al. (2013) investigated the transcriptome of *S. cerevisiae* in response to a temperature shift from 28 to 41°C to understand the adaptation of yeast cells to high temperature bioprocesses, over a period of 6 h, in batch cultures. In the initial response during the first hour, in which the cell growth was impaired, genes involved in energy metabolism, trehalose metabolism, the genes encoding molecular chaperones were most significantly induced, and the genes coding for components of translation and transcription machinery were provisionally downregulated. Biological processes related to amino acid biosynthesis, nucleotide metabolism, ion transport, and rRNA biosynthesis were found to be induced after 60 min within the cell growth permissive period. The upregulation of stress responsive genes was also observed during this period except the genes involved in trehalose metabolism. Transporters and the genes associated with purine metabolism were downregulated. The genes involved in the PKC1 pathway, also known as cell wall integrity pathway, were found to be upregulated upon heat shock and remained active for about 1 h during heat exposure.

Hughes Hallett et al. (2014) investigated the effect of the heat stress on the TORC1 pathway by exposing yeast cells to 42°C for 20 min. The response was similar to the response of yeast cells to oxidative stress; TORC1–Sch9 signaling was blocked, whereas TORC1–PP2A-branch signaling was weak or moderate.

## Cold Shock or Adaptation to Cold

Acclimatization/adaptation of yeast cells to cold was also extensively studied through perturbation experiments. First perturbation experiments to investigate the dynamic transcriptomic response of yeast cells to a temperature shift from 37 to 25°C were carried out by Gasch et al. (2000). Unlike to the response observed in a temperature shift from 25 to 37°C, the ESR genes were repressed under this condition and showed a very rapid transition to steady-state characteristics at 25°C.

Sahara et al. (2002) performed perturbation experiments to analyze the dynamic transcriptional response in wild-type yeast cells upon exposure to 10°C for 8 h. Genes involved in rRNA synthesis and the biosynthesis of ribosomal proteins were upregulated in the early phase within 30 min and in the middle phase within 2 h, respectively. General stress genes, including the genes involved in trehalose and glycogen biosynthesis, were observed to be induced in the late phase. Additionally, data indicated that cAMP–PKA pathway might have a role in the regulation of these genes. In another study aiming to investigate the adaptation of yeast cells to cold, wild-type and *MSN2/MSN4* deleted yeast cells were exposed to 10°C for varied time periods changing up to 60 h. This study has also provided additional support that ESR genes, including the genes encoding various heat shock proteins and glutathione/glutaredoxin system, were upregulated during late cold response, which is dependent on Msn2p/Msn4p. The genes involved in RNA metabolism and lipid metabolism were induced during early cold response which was Msn2/Msn4 independent (Schade et al., 2004).

The analysis of the transcriptomic response of yeast cells grown in batch cultures at 25°C and exposed to near-freezing temperature for different periods of time ranging from 6 to 48 h revealed that the genes involved in trehalose and glycogen synthesis and the genes encoding phospholipids, mannoproteins, cold shock proteins, heat shock proteins, and glutathione were upregulated for cold adaptation. The downregulation of the genes involved in protein synthesis at 4°C is in agreement with the observed delay in growth (Murata et al., 2006).

Tai et al. (2007) studied the genome-wide expression of *S. cerevisiae* cells in response to suboptimal temperatures at steady state. Yeast cells were grown in anaerobic glucose-limited and ammonium-limited chemostat cultures at a dilution rate of 0.03 h<sup>-1</sup>, at 12 and 30°C. At low temperature (12°C), transcription levels of ribosome-biogenesis genes were increased, and in contrast to batch cultures, the expression levels of 88% of ESR genes were decreased. A group of genes involved in nuclear export and ribosome biogenesis was found to be upregulated and the genes involved in carbohydrate metabolism, transport, and response to stimulus were downregulated under both nitrogen and carbon limited



conditions. This study revealed adaptational differences between the long-term exposure and a rapid shift to low temperature pointing out no need for trehalose and glycogen for the cold adaptation at steady state.

Comparative transcriptomic analysis of the effect of the low temperature (15°C), between a laboratory and a wine strain grown under anaerobic nitrogen-limited conditions in chemostats resulted in the identification of strain-specific and temperature-dependent genes. The absence of induction of the genes mediated by stress response elements implied that the GSR was repressed under 15°C in comparison to 30°C. The genes involved in trehalose metabolism and in GSR were found to be downregulated, and the genes involved in ribosome biogenesis, RNA processing were upregulated at low temperature in both strains. Integration of the transcriptome with the metabolic topology indicated that glycogen metabolism, amino acid transport, glycolipid biosynthesis, arginine biosynthesis, and allantoin metabolism were affected by the temperature. The transcript levels of the genes involved in sugar uptake and nitrogen metabolism, and transcript levels of genes related to organoleptic properties were significantly different between the two strains. The expression level of *HSF1*, encoding a heat shock transcription factor, which is active under diverse stress conditions, was lower at 15°C in both strains (Pizarro et al., 2008). García-Ríos et al. (2014) also investigated the adaptation of two different wine strains, grown in chemostat cultures at 0.028 h<sup>-1</sup>, to 15°C in order to improve the wine aroma. The integrative analysis of transcriptome with metabolome and proteome data revealed the upregulation of the sulfur assimilation pathway and glutathione biosynthesis during adaptation to cold and the response to low temperature was found to be strain specific.

The details of all these perturbation experiments in *S. cerevisiae* cells are summarized in **Table 3**. One of the common observations was that the set of the genes affected by temperature

up- or downshift was different in the early and late phases of the perturbation in batch cultures. *MSN2/MSN4*-dependent ESR genes were found to be upregulated in the late phase during temperature downshift and in the early phase during the temperature upshift. The majority of the upregulated genes were found to be downregulated upon long-term exposure to lower temperatures in chemostat experiments, where growth rate is constant. TORC1–Sch9, cAMP/PKA, and PKC1 pathways were found to be involved in the organization of the transcriptional response to heat shock and induction of the heat shock proteins was mediated by Hsf1p. High osmolarity glycerol (HOG) pathway was reported to be involved in the adaptation to cold stress that provokes changes in the membrane fluidity, which are sensed by Sln1p (Hayashi and Maeda, 2006; Panadero et al., 2006). However, the construction of quantitative models requires further studies as explained in the previous sections and remains as a challenging point.

## TRANSCRIPTOMIC RESPONSE TO SALT AND OSMOTIC SHOCK

Cells exposed to increased osmolarity, leading to water loss and cell shrinking, need to maintain their shape and turgidity. For optimal functioning of biochemical reactions appropriate concentrations of ions are required in the cytosol and organelles. HOG pathway, which is a mitogen-activated protein kinase (MAPK) signal transduction system, was reported to be the major pathway in the adaptation of yeast cells to increased osmolarity by inducing glycerol formation (Hohmann, 2009). Osmostress induction was selected as a model system by several groups to understand the regulation of gene expression by stress-activated kinases and signal transduction. The high osmolarity signaling has been reviewed by Hohmann (2009) and de Nadal and Posas (2015). The detailed description of HOG pathway and

**TABLE 3 | Perturbation experiments to monitor the transcriptomic response of *S. cerevisiae* to the changing temperatures.**

Perturbation	Experiment type	Exposure (sampling) time	Measurement technique	Reference
<b>Heat shock or adaptation to higher temperatures</b>				
From 25 to 37°C	Shake flasks	5, 15, 30, and 60 min	cDNA microarrays	Gasch et al. (2000)
From 25 to 37°C	Shake flasks	0, 15, 30, 45, 60, and 120 min	Affymetrix YE6100 gene chips	Causton et al. (2001)
25 or 39°C	NA	20 min	Agilent Yeast Oligo Microarray	Ruiz-Roig et al. (2010)
From 28 to 41°C	Batch bioreactor	0, 10, 30, 60, 120, 240, and 360 min	Affymetrix Yeast Genome S98 Array	Mensonides et al. (2013)
From 30 to 42°C	Shake flasks	20 min	Agilent Yeast Oligo Microarray (V2)	Hughes Hallett et al. (2014)
<b>Cold shock or adaptation to cold</b>				
From 37 to 25°C	Shake flasks	5, 15, 30, 45, 60, and 90 min	cDNA microarrays	Gasch et al. (2000)
From 30 to 10°C	NA	15, 30, 120, 240, and 480 min	cDNA microarrays	Sahara et al. (2002)
From 30 to 10°C	Shake flasks	10, 30, and 120 min; some cases 60 h	DNA microarrays	Schade et al. (2004)
From 25 to 4°C	Batch/time series (up to 48 h)	6, 12, 24, and 48 h	cDNA microarrays	Murata et al. (2006)
30 and 12°C	Chemostat at 0.03 h <sup>-1</sup>	Sampling at steady state	Affymetrix Yeast Genome S98 Array	Tai et al. (2007)
30 and 15°C	Chemostat at 0.05 h <sup>-1</sup>	Sampling at steady state	Affymetrix Yeast Genome 2.0 Array	Pizarro et al. (2008)
28 and 15°C	Chemostat at 0.028 h <sup>-1</sup>	Sampling at steady state	Yeast Array	García-Ríos et al. (2014)

its repressors were also extensively reviewed (Gehart et al., 2010; Saito and Posas, 2012; Engelberg et al., 2014). Several perturbation experiments were carried out to monitor the genome-wide transcriptomic response of yeast cells to the changing osmolarity in the environment.

The genome-wide analysis of the transcriptomic data of the wild-type and *HOG1* deleted yeast cells to saline stress created by the exposure to 0.4 or 0.8M NaCl for 10 or 20 min revealed that the genes involved in carbohydrate, glycerol, trehalose, and glycogen metabolism, the genes involved in the synthesis of ribosomal proteins, protein biosynthesis, and amino acid metabolism were induced after 10 min exposure to 0.4 or 0.8M NaCl. The genes associated with stress, signal transduction, and ion homeostasis were also found to be upregulated. The induction in the expression levels of the majority of genes was dependent on Hog1p. The examination of the transcriptomic response after 20 min indicated that this response was transient (Posas et al., 2000). Rep et al. (2000) analyzed the transcriptional response of the wild-type, *HOT1*, *MSN2* *MSN4*, and *HOG1* deleted strains of *S. cerevisiae* to osmotic stress created by exposing cells to 0.5 or 0.7M NaCl or 0.95M sorbitol. This study has also revealed the upregulation of the similar set of genes that is reported by Posas et al. (2000), and the induction of the majority of them were found to be Hog1p dependent. Hot1p, which is now known as the transcription factor required for the genes involved in the synthesis of glycerol, was reported to be required for the normal expression of a set of genes of HOG pathway. The authors suggested that *MSN2/MSN4* might also be regulated by Hog1p.

Analysis of the dynamic transcriptomic response of *S. cerevisiae* cells grown to mid-log phase and treated with 1M NaCl for different periods including 0, 10, 30, and 90 min revealed that the number of salinity-induced genes increases over time. Early (10 and 30 min) transcriptional response genes were found to be involved in nucleotide and amino acid metabolism, intracellular transport, protein synthesis, and destination. A few components of signaling pathways were also found upregulated in this phase. Highly expressed transcripts identified after 90 min of the treatment included salinity stress-induced genes, transporters of the major facilitator superfamily, the genes involved in the metabolism of energy reserves, nitrogen and sulfur compounds biosynthesis, and lipid, fatty acid/isoprenoid biosynthesis. The genes involved in glycerol biosynthesis (*GPD1/2*, *GPP1/2*) were observed to be upregulated at all time points (Yale and Bohnert, 2001).

The analysis of the dynamic transcriptomic response of wild-type and strains that are blocked at various points in the HOG pathway, to various concentrations of KCl and 1M sorbitol revealed that Hog1p functions during gene induction and repression, cross talk inhibition, and in governing the regulatory period. Both branches of the HOG pathway were found to be active at high osmolarity and Ssk–Sln pathway has an important role in response to modest osmolarity (O'Rourke and Herskowitz, 2004).

Fine-tuning of the response to osmotic stress at translational level was also examined by several investigators. Melamed et al. (2008) investigated the translational response of yeast cells to high osmotic stress and its correlation with the transcriptomic response. This study revealed the accumulation of non-translated RNA corresponding to a set of genes. Most of the translationally

regulated genes were found to be independent of the HOG pathway. The translational regulation of the HOG pathway-dependent genes was found to be mediated by Pub1p, and it has been suggested that the involvement of additional signaling pathways in the coordination of translational regulation. Analysis of the correlation with transcription and translation, by monitoring the affinity tagged ribosomes, indicated that changes in the transcriptome are well correlated with translation when the yeast cells were exposed to 1M sorbitol for 10 min to induce a severe osmotic stress and less correlated with the mild stress (Halbeisen and Gerber, 2009). Warringer et al. (2010) reported that the translationally regulated transcripts were dependent on Hog1p and Rck2p after hyperosmotic shock.

A yeast quantitative model of the Hog1 MAPK-dependent osmotic stress response was constructed by integrating immunoprecipitation data (ChIP-chip) with the transcriptome obtained from the analysis of single- and multiple-mutant strains (total of 31 different strains) exposed to 0.4M KCl for 20 min by Capaldi et al. (2008). This model revealed the interaction of Hog1 and Msn 2/4 pathways in information processing and regulation of gene expression in response to osmotic stress, which is context dependent. Chasman et al. (2014) have recently reported a very detailed integrative study on the pathway connectivity and the coordination of the signal in yeast cells in response to 0.7M NaCl. The transcriptomic data from 16 relevant mutants, which are carrier of the deletions in the genes known to be involved in the NaCl-induced acquired stress tolerance, was integrated with protein interaction data, and phospho-proteomic changes. This study shed light into the regulation and coordination of ESR genes and RNA Pol II was found to be key decision point in the coordination of balance between induced and repressed ESR. Cdc14p was found to be a critical integrator linking HOG and CK2 signaling, connecting to other pathways, including TORC1 and Ras/cAMP/PKA.

An integrative analysis carried out by Hughes Hallett et al. (2014), to investigate the effect of osmotic stress on the TORC1 pathway process information, revealed that the genes in the TORC1–Sch9 pathway were downregulated, and the expression levels of the genes in the TORC1–PP2A were not or little affected under osmotic conditions created with 0.4M KCl for a period of 20 min. TORC1–PP2A-branch signaling was weak, and Npr2p/Npr3p, Gtr1p/Gtr2p, and Rho1p and Snf1p were not involved in this signaling process. MAPK Hog1/p38 was found to be important in the inhibition of TORC1–Sch9 signaling, but not in other stress conditions caused by different perturbations of the system.

A comparative analysis of the transcriptomic response of two *S. cerevisiae* strains, a laboratory strain and a brewing strain, in response to high NaCl concentrations, revealed that the alterations in the expression levels of genes were larger in the laboratory strain. The response to the lower concentration of salt was rapid than that to the higher concentration in both strains. Under high NaCl concentration conditions, genes involved in carbohydrate metabolism and energy production were upregulated in both strains. Depending on the transcriptome profiles, target genes to construct a new strain with a better salt tolerance were identified, and the outcome of over-expression of those genes (*GPD1*, *ENA1*, and *CUP1*) was verified under high salinity stress (Hirasawa et al., 2006).

The details of the perturbation experiments in *S. cerevisiae* cells in response to salt and osmotic shock are summarized in **Table 4**. The cellular response of yeast cells to high osmotic perturbation was shown to be controlled by Hog1p, which also regulates the activities of Msn2p/Msn4p and TORC1 signaling pathway. Both branches of the HOG pathway were found to be active at high osmolarity, and Ssk–Sln pathway was shown to have an important role in response to modest osmolarity. A quantitative and explicit network model implicating pathway connectivity and coordination of signaling in response to osmotic stress was constructed using an integrative approach by Capaldi et al. (2008) and Chasman et al. (2014), respectively. However, the signaling dynamics and the identification of common and context-dependent features of oxidative response signaling remain to be elucidated.

## TRANSCRIPTOMIC RESPONSE TO PERTURBATIONS IN METAL ION HOMEOSTASIS

Many essential cofactors in the cell are transient metal ions, and they are functional in a range of biological processes, such as cell energetics, gene regulation, and control of free radicals (Cohen et al., 2000). Although these cations are an essential part of nutrition, they are toxic at elevated levels, causing oxidative stress, or changes in enzyme and protein function, lipid peroxidation, and DNA damage. *S. cerevisiae*, one of the most intensively studied simple eukaryote, is a good model to study how the metabolism is affected by variations in response to the availability of metal ions due to the high degree of conservation among these mechanisms concerning yeast and other higher eukaryotes. There is a substantial amount of information regarding how *S. cerevisiae* deals with these metals and metalloids and copes with them at toxic levels (Wysocki and Tamás, 2010).

Among these metals, iron and copper are of particular interest due to their ability to donate and accept electrons in vital electron

transfer reactions thus establishing themselves irreplaceable roles in many cellular processes (De Freitas et al., 2003). It also shed light on the pathophysiology of the related disorders of these processes in higher eukaryotes caused by the deficiency or the overload of these metal ions (Askwith and Kaplan, 1998).

## Response to Iron

Yeast cells respond to change in the iron availability by regulating the expression of iron transporters at transcriptional level. Aft1p was the first identified transcription factor that has a key role for the stimulation of the iron uptake systems in response to low-iron conditions (Yamaguchi-Iwai et al., 1995). Later, the genes regulated by AFT2, which is the paralog of AFT1, were identified using DNA microarrays. This study revealed that Aft2p encodes a transcription factor that has a role in the regulation of the expression in response to growth under low-iron conditions (Rutherford et al., 2001). Courel et al. (2005) designed the perturbation experiment to study the expression of genes involved in iron homeostasis by comparing the functions of Aft1p and Aft2p in regulation. Comparison of the transcriptome of the wild-type strain and corresponding AFT1 and AFT1 AFT2 deleted mutant cells, grown exponentially under iron-deficient conditions, revealed that Aft1p and its paralog Aft2p regulates the expression of genes related to iron-siderophore transport at the plasma membrane, vacuolar iron transport, and mitochondrial iron metabolism under iron-deficient conditions. Aft2p was identified to have a more specific role in the regulation of genes related to mitochondrial and vacuolar iron homeostasis, while Aft1p explicitly activates the expression of genes related to cell surface iron uptake systems.

Apart from the response observed in the iron ion transport to maintain the iron ion homeostasis, the genes involved in different metabolic pathways requiring iron-dependent enzymes were identified to be transcriptionally regulated in response to a change in the iron availability to remodel the metabolic activities for more efficient use of iron. Metabolic remodeling of yeast cells to iron deficiency was reviewed in detail by Philpott et al. (2012).

**TABLE 4 | Perturbation experiments to monitor the transcriptomic response of *S. cerevisiae* to salt and osmotic shock.**

Perturbation	Experiment type	Exposure (sampling) time	Measurement technique	Reference
0.4 or 0.8M NaCl	NA	10 and 20 min	DNA microarrays	Posas et al. (2000)
0.5 or 0.7M NaCl or 0.95M sorbitol	NA	45 min of 0.7M NaCl, 30 min of 0.5M NaCl, or 30 min of 0.95M sorbitol	Gene filters	Rep et al. (2000)
1M NaCl	NA	0, 10, 30, and 90 min	Gene filters	Yale and Bohnert (2001)
1M sorbitol or varying concentrations of KCl (0.0625, 0.125, and 0.5M)	Shake flasks	Sorbitol at 0, 5, 10, 20, 30, 40, 60, 90, 120, and 180 min KCl at 5, 10, 20, and 30 min	DNA microarrays (made in-house)	O'Rourke and Herskowitz (2004)
1M NaCl	NA	0, 10, 30, 60, 180, and 300 min	DNA microarrays	Melamed et al. (2008)
1M sorbitol	NA	10 min	Oligonucleotide microarrays	Halbeisen and Gerber (2009)
0.4M NaCl	NA	0, 2, 6, and 30 min	Yeast 6.4K microarrays	Warringer et al. (2010)
0.4M NaCl	Shake flasks	20 min	Agilent G4140A microarrays	Capaldi et al. (2008)
0.7M NaCl	Shake flasks	30 min	DNA microarrays	Chasman et al. (2014)
from 30 to 42°C	Shake flasks	20 min	Agilent Yeast Oligo Microarray (V2)	Hughes Hallett et al. (2014)
0.25, 0.5, 0.75, and 1M NaCl	Shake flasks	0, 15, 30, 45, 60, and 120 min	Yeast Gene Chip ver. 2 (DNA Chip Research Inc., Japan)	Hirasawa et al. (2006)

Shakoury-Elizeh et al. (2004) comparatively investigated the transcriptomic response of the wild-type and that of the mutant yeast cells (*AFT1* deleted or overexpressing *Aft1p*) grown in low (20  $\mu$ M), iron sufficient (100  $\mu$ M), and high (500  $\mu$ M) iron conditions, until mid-exponential phase. In addition to the identification of novel target genes of *Aft1p*, biotin uptake was reported to be upregulated under iron deficiency, while cells preferred to synthesize it when iron is abundant. Similarly, many genes involved in the synthesis and uptake of amino acids that require Fe-S proteins were reported to be regulated by iron level. Regulation of glutamate synthesis was identified to be dependent on the iron availability. Its synthesis from ammonia and  $\alpha$ -ketoglutarate was found to be catalyzed by the enzymes encoded by *GDH1* and its paralog *GDH3* under iron-deprived conditions. On the other hand, it was synthesized from glutamine and ketoglutarate under iron overloaded conditions through activity of *GLT1*, which was highly expressed. The integrative analysis of the transcriptome with metabolome data of the yeast cells grown in the presence of low (10  $\mu$ M), optimal (100  $\mu$ M), and high (330  $\mu$ M) concentration of iron for 4 h revealed that the glucose metabolism, amino acid synthesis, ergosterol, and lipid biosynthesis biological processes were all affected due to the loss in the activities of specific iron-dependent enzymes under iron deprivation. However, the amino acid homeostasis was not found to be very much affected from iron deficiency. Iron uptake systems were upregulated to preserve the activity of the iron-containing enzymes. It has been suggested that yeast cells do not have a specific machinery to forward iron ions to be used in a specific process under iron deficiency (Shakoury-Elizeh et al., 2010).

Puig et al. (2005) studied and compared the transcriptional response of yeast cells grown to the exponential phase under iron deprivation condition achieved by the addition of iron chelator or in 300  $\mu$ M  $\text{Fe}^{1+}$ -containing medium. The gene encoding fatty acid desaturase and genes related to sterol biosynthesis were upregulated, and the genes associated with TCA cycle, mitochondrial electron transport chain, heme biosynthesis, and biotin synthesis were downregulated under iron-deprived conditions. The same group further studied the role of *CTH2*, which was identified to be specifically induced under iron-deprived conditions, in iron regulon. The transcriptional response of the *CTH2* and *CTH1* *CTH2* deleted mutants to iron deprivation indicated that *Cth2p* is involved in the targeted degradation of the transcripts coding for proteins involved in multiple Fe-dependent metabolic pathways, including the TCA cycle, respiration, lipid metabolism, heme biosynthesis, and multiple Fe-S proteins. It has been suggested that metabolic remodeling in response to iron deprivation is coordinated through targeted degradation of mRNAs encoding proteins, which are involved in Fe-dependent processes and mediated by *Cth2p*. Further studies conducted by the same group (Puig et al., 2008) identified *Cth1p* having also an important role in the targeted degradation, specifically mRNAs encoding proteins involved in mitochondrial oxidative phosphorylation, while *Cth2p* preferentially involved in the targeted degradation of mRNAs of iron-containing enzymes and mRNAs associated with iron homeostasis. These two proteins were also reported to have important roles in the degradation of mRNAs involved in the transport and metabolism. An increase in the glycogen level

and activation of *Snf1p* was also observed in response to iron deficiency.

The analysis of the transcriptional response of the wild-type and iron sensitive *CCC1* deleted cells to high levels of iron under aerobic and anaerobic conditions, grown in the presence of 3 mM iron for 3 h, revealed that high iron alters the expression of the genes involved in cell-cycle progression, DNA repair, and oxidative response and iron toxicity in the *CCC1* deleted cells. The same study also revealed that iron toxicity may not be only due to oxidative damage since the increase in the transcripts indicative of oxidative damage or DNA repair in response to high iron levels was not observed in the cells under anaerobiosis. Iron sensitivity caused by the deletion of *CCC1* was reported to be suppressed by the upregulation of the genes encoding mitochondrial iron transporters *Mrs3p* or *Mrs4p* or mitochondrial pyrimidine phosphate transporter *Rim2p*. In this study, it was suggested that cells may decrease cytosolic iron levels by a mechanism that sequester iron ions into mitochondria under iron toxicity (Lin et al., 2011).

In order to investigate the role of the *Yap5p* transcription factor in the transcriptional reorganization of yeast cells, the transcriptomic response of the *YAP5* deleted cells to iron overload was examined after 20 and 60 min exposure to 2 mM  $\text{FeSO}_4$  and compared with that of the wild-type yeast cells (Pimentel et al., 2012). In addition to the alterations in the expression levels of the genes involved in iron homeostasis, this study revealed that the genes involved in ribosome biogenesis were downregulated and those of involved in stress response, protein degradation, respiration, lipid, fatty acids, and carbohydrate metabolism were upregulated under high iron-containing conditions, indicating that iron overload causes a GSR. *Yap5p* was found to be involved in the regulation *GRX4*, which regulates the nuclear localization of *Aft1p*, and the expression of *CCC1*, which is involved in iron storage, was found to be partially regulated by *Yap5p*. A schematic model to illustrate the hypothetical role of *Yap5p* under iron overload was also proposed by the authors. The regulation of iron metabolism by low and high iron sensing transcription factors (*Aft1p/Aft2p* and *Yap5p*) and post-transcriptional regulation of iron metabolism was reviewed by Outten and Albetel (2013).

## Response to Copper

The transcriptional response to low and high copper levels was reported to be regulated by transcription factors *Mac1p* and *Ace1p*, respectively (Jungmann et al., 1993). Gross et al. (2000) investigated the genome-wide transcript profiles under copper deprived and excess conditions by exposing the cells to 100  $\mu$ M  $\text{CuSO}_4$  for 30 min, to identify the targets of *Mac1p* and *Ace1p*. These experiments revealed that *Mac1p* activates the expression of *CTR1*, *CTR3*, *FRE1*, *FRE7*, *YFR055w*, and *YJL217w* under copper deficient conditions. Under copper overloaded conditions *Ace1p* induces the expression of the genes encoding metallothionein that chelates excess copper, *Cup1p* and *Crs5p*, and cytosolic copper-zinc superoxide dismutase, *Sod1p*, which detoxifies superoxide. The expression of the two genes that have roles in iron uptake system, namely *FET3*, which is required for high-affinity iron uptake, *FTR1* which forms complex with *Fet3p*, were also found to be downregulated under high copper conditions (Gross et al., 2000).



The genome-wide time-course gene expression analysis during copper starvation and copper overload in the presence of 8  $\mu\text{M}$   $\text{CuSO}_4$  for a period up to 48 h highlighted the connection between copper and iron metabolism. In response to copper deprivation, cells induce the expression of copper-independent, non-reductive iron transport genes although the global cellular iron levels did not decrease. Similar to the response observed under iron-deprived conditions, copper deficiency also lead to downregulation of the genes involved in respiration (van Bakel et al., 2005).

Cankorur-Cetinkaya et al. (2013) compared the transcriptional profiles of the wild-type and *CCC2* deleted cells under copper deficient and high or low levels of copper-containing conditions by growing yeast cells without copper or 0.04  $\mu\text{M}$  or 0.5 mM  $\text{CuSO}_4$ . *CCC2* is the human ortholog of human *ATP7A* and *ATP7B*, in which mutations are the cause of Menkes and Wilson diseases, respectively. Experimental design used in this study enabled the identification of the genes and biological processes affected from the deletion of the *CCC2* gene or the changing extracellular copper levels or the interactive effect of both factors. This study also showed the relation between copper and iron metabolisms highlighting the alteration in the transcriptional response to different level of copper availability in the absence of *CCC2*, which is involved in the copper export from the cytosol. Ribosome biogenesis and copper import were found to be downregulated in reference yeast cells in response to changes from low/deficient copper condition to high copper condition. This study revealed the processes, regulation of which under different copper levels changes depending on the presence or absence of *CCC2*. The genes involved in iron ion homeostasis, siderophore transport were identified to be upregulated in the reference strain, whereas downregulated in the absence of *CCC2* deleted cells in response to high copper levels. The iron homeostasis, siderophore transport, and  $\text{NAD}^+$  metabolism were identified to be downregulated in the deletion mutant under high copper-containing conditions and upregulated under copper-deficient conditions when compared to the reference strain under same conditions. The amino acid metabolism, specifically arginine metabolic process, was also identified to be altered by the interactive effect of both perturbations, and these findings were also supported by the metabolomic analysis. The transcription factors, around which most transcriptomic changes occur, were also identified in this study through integrative analysis of transcriptome and regulome. This integrative analysis indicated that the genes targeted by general oxidative stress inducers, namely *Sko1p*, *Skn7p*, *Cin5p*, *Yap1p*, and *Yap6p*, were mostly affected from both perturbations and their interactions.

## Response to Other Metal Ions

The transcriptional response of yeast to zinc deficiency was studied in glucose- and ammonium-limited chemostat cultures aerobically and anaerobically (De Nicola et al., 2007). Zinc-specific *Zap1p* regulon, a central transcription factor that is active in response to zinc alterations, was identified to be the regulator of genes involved in carbohydrate storage metabolism. It was found that oxygen and Zn availability affected a large number of genes, implying a more significant role of Zn in mitochondrial processes. Wu et al. (2008) used transcriptome profiling of the

wild-type and mutant cells, *ZAP1* deleted and cells containing plasmids encoding constitutive allele of *Zap1*, to detect the target genes of *Zap1p*. Yeast cells were exposed to various concentrations of zinc changing from 3 to 300  $\mu\text{M}$  for various periods of time changing from 0.5 to 8 h. In addition to their previous study (Lyons et al., 2000), in which 46 genes were identified to be potential target genes; in this study, they further investigated the role of *Zap1p* and identified numerous new targets of *Zap1*-mediated regulation. The transcriptomic response of the *Zap1* target genes was shown to have dependency to the level of zinc in the medium. This study revealed that cells induce the genes involved in zinc uptake to maintain the zinc homeostasis in response to a mild zinc deficiency. On the other hand, the genes involved in maintaining secretory pathway and cell wall function, and stress responses were regulated at transcriptional level in response to a severe zinc deficiency. It has been suggested that these group of genes are mainly involved in the adaptation to zinc deficiency.

Jin et al. (2008) investigated the genome-wide response of yeast cells to a 2 h of exposure to equitoxic concentrations of  $\text{Zn}^{2+}$ ,  $\text{Cd}^{2+}$ ,  $\text{Hg}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Ag}^+$ ,  $\text{Cr}^{6+}$ , and  $\text{As}^{3+}$  and complemented this study with deletome. Principal component analysis (PCA) and hierarchical clustering of global transcriptomic response indicated that the response was specific for each metal and the response to different concentrations of the same metal were closely related but different. A group of genes, called common metal responsive (CMR) genes, were found to be commonly affected by all metals. The genes involved in metal transport and homeostasis, detoxification of ROS, carbohydrate metabolism, including glycolysis, oxidative phosphorylation, and alcohol metabolism, polyamine transport, and transcription were upregulated. The genes involved in polysaccharide metabolism, G-protein signaling, protein targeting, and transport were downregulated. Some evolutionarily conserved, signal transduction pathways, including cAMP-dependent PKA, protein kinase CK2, and MAPK, were found to be involved in the regulation of responses to the exposure to metals. *Msn2p*/*Msn4p* was observed to regulate 10% of the differentially expressed genes, among the 14 perturbation conditions.

A short-term effect of the moderate amounts of metals on *S. cerevisiae* cells was investigated (Hosiner et al., 2014). Analysis of the transcriptomic data after 30 min acquaintance to  $\text{Ag}^+$ ,  $\text{Al}^{3+}$ ,  $\text{As}^{3+}$ ,  $\text{Cd}^{2+}$ ,  $\text{Co}^{2+}$ ,  $\text{Hg}^{2+}$ ,  $\text{Mn}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{V}^+$ , and  $\text{Zn}^{2+}$ , metals relevant to human health indicated that the metal-specific oxidative defense, including glutathione/thioredoxin and metallothionein systems, and protein degradation processes, including vacuolar protein degradation, proteosomal proteolysis, chaperone complex activities, and *Sec19*, which regulates vesicle traffic in secretory pathways, were activated in response to all these metal ions. The genes involved in ribosome biogenesis were observed to be downregulated after a short-term exposure to metals. The potential regulators effective under these conditions were predicted *via* a statistical tool, and the largest group covered the transcription factors *Yap1p*, *Msn2p*, *Msn4p*, *Yap7p*, and *Cad1p*, the latter two being the homologs of *Yap1p*. A comparative analysis with the results obtained after 2 h exposure to metals (Jin et al., 2008) showed the induction of the genes involved in protein synthesis and a repression of the genes associated with the metal detoxification process.

The details of all these perturbation experiments in *S. cerevisiae* cells are summarized in **Table 5**. All these studies indicated that yeast cells reorganize their transcriptional response and their metabolism upon exposure or deficiency of transition metals.

Since the iron deficiency is the most common, worldwide nutritional disorder, first perturbation experiments related to iron were designed to understand the effect of the iron deficiency on the reorganization and the regulation of transcriptomic response. These studies resulted in the identification of the target genes controlling iron transport and homeostasis as well as the post-transcriptional regulation of the iron-dependent metabolic pathways under iron deficiency. Aft1p, Aft2p, Yap5p, and Snf1p were found to be the major transcription factors in regulation of the cellular reorganization under iron deprivation.

Although a number of experiments were also carried out to understand the effect of the iron overload, the results of these experiments are not comparable due to the use of different iron concentrations and exposure times. A set of carefully designed perturbation experiments should be planned to understand the effect of the iron overload and to reveal the signaling pathway(s) underlying the response to iron overload and deficiency.

Early copper-related genome-wide studies are concentrated on the identification of target genes of two transcriptional factors in response to copper deficiency and copper overload. However,

these pioneering studies are not easily comparable due to the differences in experimental design as in the case of iron.

The perturbation experiments carried out by exposing yeast cells to low or high levels of different metal ions for a defined period of time revealed a detailed picture for the organization of the transcriptional response to metals, including copper. Evolutionarily conserved signaling pathways, including cAMP-dependent PKA, protein kinase that coordinate the post-translational regulation of the proteins involved in the glycolysis and gluconeogenesis and the transcriptional regulation of the synthesis of ribosomal protein, CK2 that is involved in ribosome biosynthesis, and MAPK that is involved in apoptosis, differentiation, and stress response were identified to participate in the organization of the response to metals. Msn2p/Msn4p, Yap1p, and its homologs Yap7p and Cad1p were found to be involved in the regulation of a large number of metal responsive genes. However, a quantitative model implicating these results is still missing.

## CONCLUSION AND PERSPECTIVE

Yeast cells may encounter many kinds of environmental perturbations during growth and fermentation, including accumulation of ethanol, weak acids, heat, low pH, ROS, nutrient limitation, and osmotic changes imposed by high concentrations of sugars

**TABLE 5 | Perturbation experiments to monitor the transcriptomic response of *S. cerevisiae* to changes in metal ion homeostasis.**

Perturbation	Experiment type	Exposure (sampling) time	Measurement technique	Reference
<b>Response to iron</b>				
Deletion or overexpression of AFT2	Shake flasks	Sampled at OD = 0.4	DNA microarrays	Rutherford et al. (2001)
<i>aft1Δ</i> and <i>aft1Δ/aft2Δ</i> cells grown in iron depletion	NA	Sampled at log phase	DNA microarrays	Courel et al. (2005)
Deletion or overexpression of AFT1 under 20, 100, or 500 μM iron-containing conditions	NA	Sampled at mid-log phase	cDNA microarrays	Shakoury-Elizeh et al. (2004)
at mid-exp. phase switch from 100 to 10, or 300 μM iron-containing conditions	NA	4 h	Affymetrix Yeast Genome 2.0 Array	Shakoury-Elizeh et al. (2010)
No iron or 300 μM iron-containing conditions	NA	Sampled at log phase	DNA microarrays	Puig et al. (2005)
<i>cth1Δ/cth2Δ</i> cells expressing physiological levels of <i>CTH1</i> , <i>CTH2</i> , or vector alone under iron-deficient conditions	NA	Sampled at log phase	Affymetrix Yeast Genome S98 Array	Puig et al. (2008)
<i>ccc1Δ</i> exposed to 3 mM of iron under aerobic and anaerobic conditions	NA	3 h	Affymetrix microarray	Lin et al. (2011)
2 mM iron	Shake flasks	20 and 60 min	UAVR Yeast 19.2K v1	Pimentel et al. (2012)
<b>Response to copper</b>				
Wild-type versus constitutively active Mac1 at mid-exp. phase switch to no copper, or 100 μM copper containing conditions	NA	60 min of no copper 30 min of 100 μM CuSO <sub>4</sub>	DNA microarrays	Gross et al. (2000)
At mid-exp. phase switch to no copper, or 8 μM copper containing conditions	NA	0.5, 1, 2, 3, 4, 24, and 48 h	A-UMCU-4 – UMC Utrecht <i>S. cerevisiae</i> 16K array, version 1.2	van Bakel et al. (2005)
Wild-type versus <i>ccc2Δ/ccc2Δ</i>	Batch bioreactor	Sampled at mid-log phase	Affymetrix Yeast Genome 2.0 Array	Cankorur-Cetinkaya et al. (2013)
No copper, 0.04 μM or 0.5 mM copper containing conditions				
<b>Other metal ions</b>				
Zinc limitation at aerobic and anaerobic conditions	Chemostat at 0.1 h <sup>-1</sup>	Sampled at steady state	Affymetrix Yeast Genome S98 Array	De Nicola et al. (2007)
Zinc replete (various Zn concentrations) and zinc limiting	NA	At log phase or transient: 0.5, 1, 2, 4, and 8 h	DNA microarray	Wu et al. (2008)
Equitoxic concentrations of Zn <sup>2+</sup> , Cd <sup>2+</sup> , Hg <sup>2+</sup> , Cu <sup>2+</sup> , Ag <sup>+</sup> , Cr <sup>6+</sup> , and As <sup>3+</sup>	NA	2 h	DNA microarray	Jin et al. (2008)
Ag <sup>+</sup> , Al <sup>3+</sup> , As <sup>3+</sup> , Cd <sup>2+</sup> , Co <sup>2+</sup> , Hg <sup>2+</sup> , Mn <sup>2+</sup> , Ni <sup>2+</sup> , V <sup>+</sup> , or Zn <sup>2+</sup>	Shake flasks	30 min	Microarrays obtained from Microarray Centre Toronto	Hosiner et al. (2014)

(Teixeira et al., 2011; Zhao et al., 2015). In most cases, cells have to cope simultaneously or successively with numerous aspects. Studying these responses is important for optimizing industrial process conditions and designing more robust overproduction strains using a rational design strategy and genetic engineering techniques. In addition to that, such perturbation/response studies are important due to their contribution to our fundamental understanding of microbial metabolism and to unravel drug-disease relationships. Here, we summarized findings on the yeast transcriptome in response to nutritional, osmotic, oxidative, temperature, and transient metal ion perturbations, as such not aimed for an exhaustive literature survey.

Transcriptomic analysis of *S. cerevisiae* subjected to different types of stress-causing conditions helped us to identify the genes with altered expression as a result of a perturbation and the use of deleted strains for the known transcription factors led to the determination of the affected genes and the up- or downregulated biological processes. However, studies carried out in the first decade since the application of the high-throughput techniques have discrepancies due to the differences in experimental design including the type of the equipment or microarrays, fermentation conditions and mode, selection of control, duration of the exposure to stress-causing agents, and the amount of stress-causing agent. The first perturbation experiments were carried out in batch. Chemostat fermentations were preferred after the elucidation of the relationship between the response and the growth rate. The statistical and bioinformatics tools used to analyze the results are also important features that make the comparative analysis very difficult.

In the last decade, carefully designed perturbation experiments considering prior genetic data and the integration of transcriptome with metabolome, proteome, phospho-proteome, and interactome revealed the presence of some shared signaling pathways, which are activated in response to several environmental perturbations and also the condition dependent response of *S. cerevisiae* to specific perturbations. Ras/cAMP/PKA signaling pathway, which is involved in cell growth and response to nutrients and stress, and TORC1, which is the main regulator of growth and metabolism in all eukaryotic cells, were found to be common in response to all perturbations considered in this review. Sch9-branch signaling and PP2A-branch signaling of TORC1 were reported to be inhibited under glucose starvation, osmotic stress, oxidative stress, and heat stress in yeast cells. Different types of nitrogen starvation and rapamycin were observed to lead to the activation of PP2A-branch signaling of TOR kinase (Hughes Hallett et al., 2014). Snf1p, AMP-activated serine/threonine protein kinase, which is involved in the transcription of glucose repressed genes in yeast, is also a common regulator in organizing the response of yeast cells to several perturbations. In addition to the common signaling pathways, ESR genes, primarily controlled by the transcription factors Msn2p and Msn4p, which are known to be regulated by Ras/cAMP/PKA pathway, and/or by TORC1, are also induced in response to various environmental perturbations. However, the role of the Msn2p and Msn4p for the acquisition of stress tolerance has remained rather obscure. While Berry and Gasch (2008) suggested a significant role of Msn2p and Msn4p in the gain of stress tolerance, Zakrzewska et al. (2011) claimed

that these two transcription factors are not functional for the acquisition of severe stress tolerance, but decline in the growth rate is critical for the gain of tolerance. However, the molecular mechanism leading to the development of tolerance, which is of extreme importance for biotechnological applications, requires further examination.

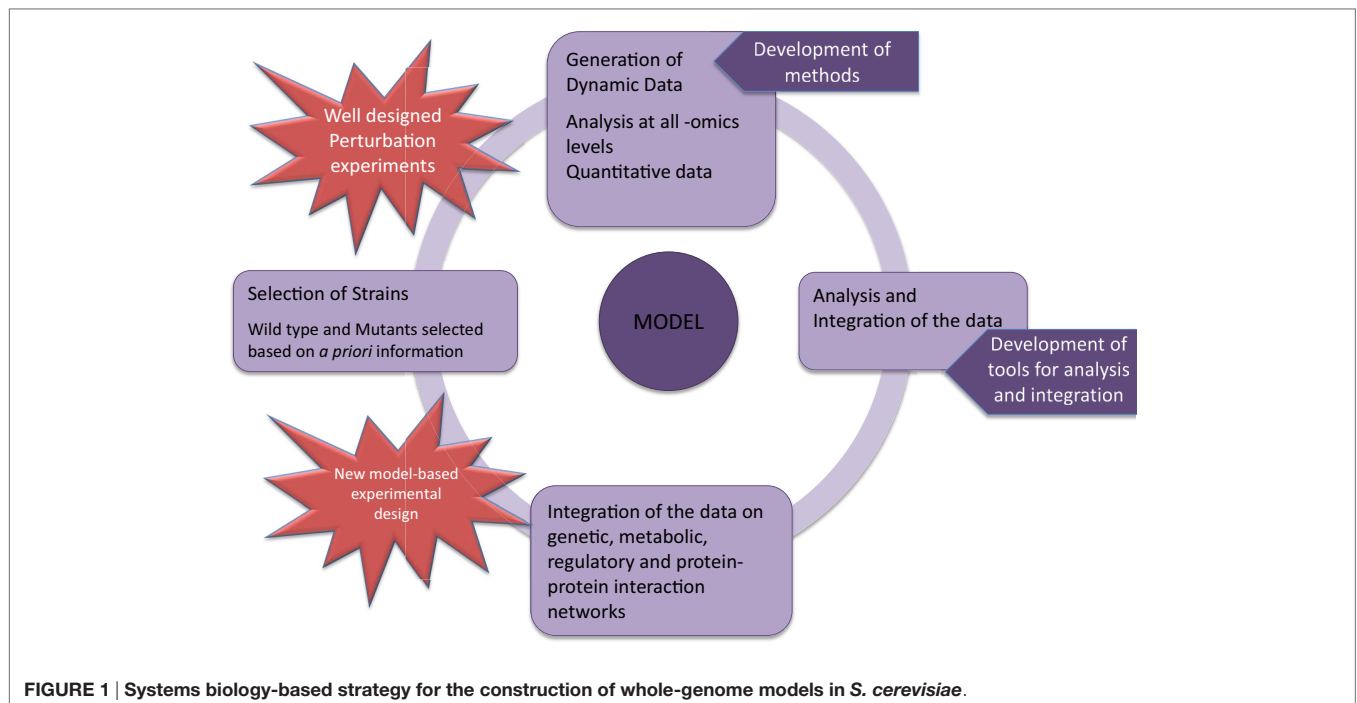
The major goal of system biology is to construct quantitative whole life models for the prediction of cellular response to the changing environment. When we overview the results of a number of studies carried out in the last 15 years, which are focused to elucidate the response to an induced stress using one or several omics technologies, it seems that we are far away to reach this aim at present. However, it should be noted that the efforts, especially on the development of integrative systems biology approaches complemented with accumulated and shared information, continue and guarantee the success in the near future.

## Future Prospects

Yeast cells have established a range of mechanisms in response to environmental and genetic perturbations, in order to adopt themselves to the new condition. This is attained by not only changes in the gene expression levels and protein regulation but also mRNA stability, non-covalent binding of allosteric effectors and post-translational modifications of enzymes are involved (Tripodi et al., 2015). Therefore, despite the large number of perturbation experiments carried out to date, understanding the complete regulatory and signaling networks in response to perturbations requires further carefully designed perturbation experiments complemented with integrative analysis of -omics and new computational approaches (see **Figure 1**). It should be noted that strain selection and selection of the mutants, which will be incorporated into the study should be carefully planned by incorporating prior genetic knowledge. The complementation of environmental perturbations with genetic perturbations using appropriate mutants will help to characterize the details and coordination of the underlying signaling and regulatory events.

Most of the reviewed articles here used microarray technology for studying transcriptome. Since the RNA-seq technology has been available for several years now, the number of studies employing this technology is expected to accumulate and play a major role in future research. The properties of RNA-seq such as providing more precise measurement of transcript levels, and excluding limitations such as non-specific hybridization and possible signal saturation due to high abundance of transcripts makes it a valuable tool (Wang et al., 2009; Nookaew et al., 2012).

Most of the studies carried out in the past on the yeast transcriptome comprise measurements at single-type point (static). Time-series transcriptome data complemented with quantitative proteome and phospho-proteome reflect the dynamic regulation of gene expression, thus incorporation of such data improves the validity of a quantitative model, derived to predict the activities of genes under a particular condition. The design of new perturbation experiments to monitor dynamic changes at all -omics levels including interactome and metabolome and the development of new computational tools to analyze and integrate this data will possibly facilitate to reveal the underlying regulatory and



signaling events and will immensely contribute to improve the predictive capability of models in the near future.

It was not within the scope of this review, but it is noteworthy that the yeast response mechanisms to cope with the presence of a foreign compound, i.e., drug, is an attractive research field to discover drug targets or understand the mechanism of action. Studies on the deletion or overexpressing mutants are considered very important in the understanding of molecular basis of diseases and in finding novel drug targets. For example, the MAP kinase Hog1 is the yeast ortholog of mammalian p38, important for embryonic development and cancer progression (Bradham and McClay, 2006). Consequently, yeast will continue to be one of the major model organisms for systems biology approaches

and specifically designed perturbation experiments will help to develop whole life models.

## AUTHOR CONTRIBUTIONS

All authors participated equally in the preparation of this contribution, have read, and approved the final manuscript.

## ACKNOWLEDGMENTS

We gratefully acknowledge the funding from The Scientific and Technological Research Council of Turkey (TUBITAK) through project no. 114C062.

## REFERENCES

- Askwith, C., and Kaplan, J. (1998). Iron and copper transport in yeast and its relevance to human disease. *Trends Biochem. Sci.* 23, 135–138. doi:10.1016/S0968-0004(98)01192-X
- Bar-Joseph, Z., Gitter, A., and Simon, I. (2012). Studying and modelling dynamic biological processes using time-series gene expression data. *Nat. Rev. Genet.* 13, 552–564. doi:10.1038/nrg3244
- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: archive for functional genomics data sets – update. *Nucleic Acids Res.* 41, D991–D995. doi:10.1093/nar/gks1193
- Berry, D. B., and Gasch, A. P. (2008). Stress-activated genomic expression changes serve a preparative role for impending stress in yeast. *Mol. Biol. Cell* 19, 4580–4587. doi:10.1091/mbc.E07-07-0680
- Bisschops, M. M., Zwartjens, P., Keuter, S. G., Pronk, J. T., and Daran-Lapujade, P. (2014). To divide or not to divide: a key role of Rim15 in calorie-restricted yeast cultures. *Biochim. Biophys. Acta* 1843, 1020–1030. doi:10.1016/j.bbamcr.2014.01.026
- Boer, V. M., Amini, S., and Botstein, D. (2008). Influence of genotype and nutrition on survival and metabolism of starving yeast. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6930–6935. doi:10.1073/pnas.0802601105
- Boer, V. M., Crutchfield, C. A., Bradley, P. H., Botstein, D., and Rabinowitz, J. D. (2010). Growth-limiting intracellular metabolites in yeast growing under diverse nutrient limitations. *Mol. Biol. Cell* 21, 198–211. doi:10.1091/mbc.E09-07-0597
- Boer, V. M., de Winde, J. H., Pronk, J. T., and Piper, M. D. (2003). The genome-wide transcriptional responses of *Saccharomyces cerevisiae* grown on glucose in aerobic chemostat cultures limited for carbon, nitrogen, phosphorus, or sulfur. *J. Biol. Chem.* 278, 3265–3274. doi:10.1074/jbc.M209759200
- Boone, C. (2014). Yeast systems biology: our best shot at modeling a cell. *Genetics* 198, 435–437. doi:10.1534/genetics.114.169128
- Bradham, C., and McClay, D. R. (2006). p38 MAPK in development and cancer. *Cell Cycle* 5, 824–828. doi:10.4161/cc.5.8.2685
- Bradley, P. H., Brauer, M. J., Rabinowitz, J. D., and Troyanskaya, O. G. (2009). Coordinated concentration changes of transcripts and metabolites in *Saccharomyces cerevisiae*. *PLoS Comput. Biol.* 5:e1000270. doi:10.1371/journal.pcbi.1000270
- Brauer, M. J., Huttenhower, C., Airolidi, E. M., Rosenstein, R., Matese, J. C., Gresham, D., et al. (2008). Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. *Mol. Biol. Cell* 19, 352–367. doi:10.1091/mbc.E07-08-0779
- Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., et al. (2001). Minimum information about a microarray experiment



- (MIAME)-toward standards for microarray data. *Nat. Genet.* 29, 365–371. doi:10.1038/ng1201-365
- Cankorur-Cetinkaya, A., Eraslan, S., and Kirdar, B. (2013). Transcriptional remodeling in response to changing copper levels in the Wilson and Menkes disease model of *Saccharomyces cerevisiae*. *Mol. Biosyst.* 9, 2889–2908. doi:10.1039/c3mb70276f
- Capaldi, A. P., Kaplan, T., Liu, Y., Habib, N., Regev, A., Friedman, N., et al. (2008). Structure and function of a transcriptional network activated by the MAPK Hog1. *Nat. Genet.* 40, 1300–1306. doi:10.1038/ng.235
- Castrillo, J. I., Zeef, L. A., Hoyle, D. C., Zhang, N., Hayes, A., Gardner, D. C., et al. (2007). Growth control of the eukaryote cell: a systems biology study in yeast. *J. Biol.* 6, 4. doi:10.1186/jbiol54
- Causton, H. C., Ren, B., Koh, S. S., Harbison, C. T., Kanin, E., Jennings, E. G., et al. (2001). Remodeling of yeast genome expression in response to environmental changes. *Mol. Biol. Cell* 12, 323–337. doi:10.1091/mbc.01-10-0472
- Chasman, D., Ho, Y. H., Berry, D. B., Nemec, C. M., MacGilvray, M. E., Hose, J., et al. (2014). Pathway connectivity and signaling coordination in the yeast stress-activated signaling network. *Mol. Syst. Biol.* 10, 759. doi:10.15252/msb.20145120
- Cohen, A., Nelson, H., and Nelson, N. (2000). *Metal-Ion Transporters – From Yeast to Human Diseases*. Austin, TX: Landes Bioscience. Madame Curie Bioscience Database [Internet].
- Cohen, B. A., Pilpel, Y., Mitra, R. D., and Church, G. M. (2002). Discrimination between paralogs using microarray analysis: application to the Yap1p and Yap2p transcriptional networks. *Mol. Biol. Cell* 13, 1608–1614. doi:10.1091/mbc.01-10-0472
- Conrad, M., Schothorst, J., Kankipati, H. N., Van Zeebroeck, G., Rubio-Teixeira, M., and Thevelein, J. M. (2014). Nutrient sensing and signaling in the yeast *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* 38, 254–299. doi:10.1111/1574-6976.12065
- Courel, M., Lallet, S., Camadro, J. M., and Blaiseau, P. L. (2005). Direct activation of genes involved in intracellular iron use by the yeast iron-responsive transcription factor Aft2 without its paralog Aft1. *Mol. Cell. Biol.* 25, 6760–6771. doi:10.1128/MCB.25.15.6760-6771.2005
- De Freitas, J., Wintz, H., Kim, J. H., Poynton, H., Fox, T., and Vulpe, C. (2003). Yeast, a model organism for iron and copper metabolism studies. *Biomaterials* 16, 185–197. doi:10.1023/A:1020771000746
- de Nadal, E., and Posas, F. (2015). Osmostress-induced gene expression – a model to understand how stress-activated protein kinases (SAPKs) regulate transcription. *FEBS J.* 282, 3275–3285. doi:10.1111/febs.13323
- De Nicola, R., Hazelwood, L. A., De Hulster, E. A., Walsh, M. C., Knijnenburg, T. A., Reinders, M. J., et al. (2007). Physiological and transcriptional responses of *Saccharomyces cerevisiae* to zinc limitation in chemostat cultures. *Appl. Environ. Microbiol.* 73, 7680–7692. doi:10.1128/AEM.01445-07
- DeRisi, J. L., Iyer, V. R., and Brown, P. O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278, 680–686. doi:10.1126/science.278.5338.680
- Dikicioglu, D., Dunn, W. B., Kell, D. B., Kirdar, B., and Oliver, S. G. (2012). Short- and long-term dynamic responses of the metabolic network and gene expression in yeast to a transient change in the nutrient environment. *Mol. Biosyst.* 8, 1760–1774. doi:10.1039/c2mb05443d
- Dikicioglu, D., Karabekmez, E., Rash, B., Pir, P., Kirdar, B., and Oliver, S. G. (2011). How yeast re-programmes its transcriptional profile in response to different nutrient impulses. *BMC Syst. Biol.* 5:148. doi:10.1186/1752-0509-5-148
- Engelberg, D., Perlman, R., and Levitzki, A. (2014). Transmembrane signaling in *Saccharomyces cerevisiae* as a model for signaling in metazoans: state of the art after 25 years. *Cell. Signal.* 26, 2865–2878. doi:10.1016/j.cellsig.2014.09.003
- Farrugia, G., and Balzan, R. (2012). Oxidative stress and programmed cell death in yeast. *Front. Oncol.* 2:64. doi:10.3389/fonc.2012.00064
- Fazio, A., Jewett, M. C., Daran-Lapujade, P., Mustacchi, R., Usaite, R., Pronk, J. T., et al. (2008). Transcription factor control of growth rate dependent genes in *Saccharomyces cerevisiae*: a three factor design. *BMC Genomics* 9:341. doi:10.1186/1471-2164-9-341
- García-Ríos, E., López-Malo, M., and Guillaumon, J. M. (2014). Global phenotypic and genomic comparison of two *Saccharomyces cerevisiae* wine strains reveals a novel role of the sulfur assimilation pathway in adaptation at low temperature fermentations. *BMC Genomics* 15:1059. doi:10.1186/1471-2164-15-1059
- Gasch, A. P. (2003). “The environmental stress response: a common yeast response to environmental stresses,” in *Yeast Stress Responses*, Vol. 1, eds Hohmann S. and Mager W. H. (Heidelberg: Springer-Verlag), 11–70.
- Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., et al. (2000). Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* 11, 4241–4257. doi:10.1091/mbc.11.12.4241
- Gehart, H., Kumpf, S., Ittner, A., and Ricci, R. (2010). MAPK signalling in cellular metabolism: stress or wellness? *EMBO Rep.* 11, 834–840. doi:10.1038/embor.2010.160
- Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., et al. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387–391. doi:10.1038/nature00935
- Giaever, G., and Nislow, C. (2014). The yeast deletion collection: a decade of functional genomics. *Genetics* 197, 451–465. doi:10.1534/genetics.114.161620
- Gross, C., Kelleher, M., Iyer, V. R., Brown, P. O., and Winge, D. R. (2000). Identification of the copper regulon in *Saccharomyces cerevisiae* by DNA microarrays. *J. Biol. Chem.* 275, 32310–32316. doi:10.1074/jbc.M005946200
- Halbeisen, R. E., and Gerber, A. P. (2009). Stress-dependent coordination of transcriptome and translome in yeast. *PLoS Biol.* 7:e1000105. doi:10.1371/journal.pbio.1000105
- Hayashi, M., and Maeda, T. (2006). Activation of the HOG pathway upon cold stress in *Saccharomyces cerevisiae*. *J. Biochem.* 139, 797–803. doi:10.1093/jb/mvj089
- Hirasawa, T., Nakakura, Y., Yoshikawa, K., Ashitani, K., Nagahisa, K., Furusawa, C., et al. (2006). Comparative analysis of transcriptional responses to saline stress in the laboratory and brewing strains of *Saccharomyces cerevisiae* with DNA microarray. *Appl. Microbiol. Biotechnol.* 70, 346–357. doi:10.1007/s00253-005-0192-6
- Ho, Y. H., and Gasch, A. P. (2015). Exploiting the yeast stress-activated signaling network to inform on stress biology and disease signaling. *Curr. Genet.* 61, 503–511. doi:10.1007/s00294-015-0491-0
- Hohmann, S. (2009). Control of high osmolarity signalling in the yeast *Saccharomyces cerevisiae*. *FEBS Lett.* 583, 4025–4029. doi:10.1016/j.febslet.2009.10.069
- Hosiner, D., Gerber, S., Lichtenberg-Fraté, H., Glaser, W., Schüller, C., and Klipp, E. (2014). Impact of acute metal stress in *Saccharomyces cerevisiae*. *PLoS ONE* 9:e83330. doi:10.1371/journal.pone.0083330
- Hughes Hallett, J. E., Luo, X., and Capaldi, A. P. (2014). State transitions in the TORC1 signaling pathway and information processing in *Saccharomyces cerevisiae*. *Genetics* 198, 773–786. doi:10.1534/genetics.114.168369
- Jin, Y. H., Dunlap, P. E., McBride, S. J., Al-Refai, H., Bushel, P. R., and Freedman, J. H. (2008). Global transcriptome and deletome profiles of yeast exposed to transition metals. *PLoS Genet.* 4:e1000053. doi:10.1371/journal.pgen.1000053
- Jungmann, J., Reins, H. A., Lee, J., Romeo, A., Hassett, R., Kosman, D., et al. (1993). MAC1, a nuclear regulatory protein related to Cu-dependent transcription factors is involved in Cu/Fe utilization and stress resistance in yeast. *EMBO J.* 12, 5051–5056.
- Kershaw, C. J., Costello, J. L., Castelli, L. M., Talavera, D., Rowe, W., Sims, P. F., et al. (2015). The yeast La related protein Slf1p is a key activator of translation during the oxidative stress response. *PLoS Genet.* 11:e1004903. doi:10.1371/journal.pgen.1004903
- Kolesnikov, N., Hastings, E., Keays, M., Melnichuk, O., Tang, Y. A., Williams, E., et al. (2015). ArrayExpress update – simplifying data submissions. *Nucleic Acids Res.* 43, D1113–D1116. doi:10.1093/nar/gku1057
- Kresnowati, M. T., van Winden, W. A., Almering, M. J., ten Pierick, A., Ras, C., Knijnenburg, T. A., et al. (2006). When transcriptome meets metabolome: fast cellular responses of yeast to sudden relief of glucose limitation. *Mol. Syst. Biol.* 2, 49. doi:10.1038/msb4100083
- Lin, H., Li, L., Jia, X., Ward, D. M., and Kaplan, J. (2011). Genetic and biochemical analysis of high iron toxicity in yeast: iron toxicity is due to the accumulation of cytosolic iron and occurs under both aerobic and anaerobic conditions. *J. Biol. Chem.* 286, 3851–3862. doi:10.1074/jbc.M110.190959
- Livas, D., Almering, M. J., Daran, J. M., Pronk, J. T., and Gancedo, J. M. (2011). Transcriptional responses to glucose in *Saccharomyces cerevisiae* strains lacking a functional protein kinase A. *BMC Genomics* 12:405. doi:10.1186/1471-2164-12-405
- Loewith, R., and Hall, M. N. (2011). Target of rapamycin (TOR) in nutrient signaling and growth control. *Genetics* 189, 1177–1201. doi:10.1534/genetics.111.133363
- Lyons, T. J., Gasch, A. P., Gaither, L. A., Botstein, D., Brown, P. O., and Eide, D. J. (2000). Genome-wide characterization of the Zap1p zinc-responsive regulon in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 97, 7957–7962. doi:10.1073/pnas.97.14.7957
- Melamed, D., Pnueli, L., and Arava, Y. (2008). Yeast translational response to high salinity: global analysis reveals regulation at multiple levels. *RNA* 14, 1337–1351. doi:10.1261/rna.864908

- Mensonides, F., Hellingwerf, K., Teixeira de Mattos, M., and Brul, S. (2013). Multiphasic adaptation of the transcriptome of *Saccharomyces cerevisiae* to heat stress. *Food Res. Intern.* 54, 1103–1112. doi:10.1016/j.foodres.2012.12.042
- Mitchell, A., Romano, G. H., Groisman, B., Yona, A., Dekel, E., Kupiec, M., et al. (2009). Adaptive prediction of environmental changes by microorganisms. *Nature* 460, 220–224. doi:10.1038/nature08112
- Morano, K. A., Grant, C. M., and Moye-Rowley, W. S. (2012). The response to heat shock and oxidative stress in *Saccharomyces cerevisiae*. *Genetics* 190, 1157–1195. doi:10.1534/genetics.111.128033
- Murata, Y., Homma, T., Kitagawa, E., Momose, Y., Sato, M. S., Odani, M., et al. (2006). Genome-wide expression analysis of yeast response during exposure to 4 degrees C. *Extremophiles* 10, 117–128. doi:10.1007/s00792-005-0480-1
- Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., et al. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* 320, 1344–1349. doi:10.1126/science.1158441
- Nookaew, I., Papini, M., Pornputtapong, N., Scalcinati, G., Fagerberg, L., Uhlén, M., et al. (2012). A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 40, 10084–10097. doi:10.1093/nar/gks804
- Oliveira, A. P., Dimopoulos, S., Busetto, A. G., Christen, S., Dechant, R., Falter, L., et al. (2015). Inferring causal metabolic signals that regulate the dynamic TORC1-dependent transcriptome. *Mol. Syst. Biol.* 11, 802. doi:10.15252/msb.20145475
- O'Rourke, S. M., and Herskowitz, I. (2004). Unique and redundant roles for HOG MAPK pathway components as revealed by whole-genome expression analysis. *Mol. Biol. Cell* 15, 532–542. doi:10.1091/mbc.E03-07-0521
- Outten, C. E., and Albetel, A. N. (2013). Iron sensing and regulation in *Saccharomyces cerevisiae*: ironing out the mechanistic details. *Curr. Opin. Microbiol.* 16, 662–668. doi:10.1016/j.mib.2013.07.020
- Panadero, J., Pallotti, C., Rodríguez-Vargas, S., Randez-Gil, F., and Prieto, J. A. (2006). A downshift in temperature activates the high osmolarity glycerol (HOG) pathway, which determines freeze tolerance in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 281, 4638–4645. doi:10.1074/jbc.M512736200
- Philpott, C. C., Leidgens, S., and Frey, A. G. (2012). Metabolic remodeling in iron-deficient fungi. *Biochim. Biophys. Acta* 1823, 1509–1520. doi:10.1016/j.bbamcr.2012.01.012
- Pimentel, C., Vicente, C., Menezes, R. A., Caetano, S., Carreto, L., and Rodrigues-Pousada, C. (2012). The role of the Yap5 transcription factor in remodeling gene expression in response to Fe bioavailability. *PLoS ONE* 7:e37434. doi:10.1371/journal.pone.0037434
- Pizarro, F. J., Jewett, M. C., Nielsen, J., and Agosin, E. (2008). Growth temperature exerts differential physiological and transcriptional responses in laboratory and wine strains of *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* 74, 6358–6368. doi:10.1128/AEM.00602-08
- Posas, F., Chambers, J. R., Heyman, J. A., Hoeffler, J. P., de Nadal, E., and Ariño, J. (2000). The transcriptional response of yeast to saline stress. *J. Biol. Chem.* 275, 17249–17255. doi:10.1074/jbc.M910016199
- Puig, S., Askeland, E., and Thiele, D. J. (2005). Coordinated remodeling of cellular metabolism during iron deficiency through targeted mRNA degradation. *Cell* 120, 99–110. doi:10.1016/j.cell.2004.11.032
- Puig, S., Vergara, S. V., and Thiele, D. J. (2008). Cooperation of two mRNA-binding proteins drives metabolic adaptation to iron deficiency. *Cell Metab.* 7, 555–564. doi:10.1016/j.cmet.2008.04.010
- Regenberg, B., Grotkjaer, T., Winther, O., Fausbøll, A., Akesson, M., Bro, C., et al. (2006). Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in *Saccharomyces cerevisiae*. *Genome Biol.* 7, R107. doi:10.1186/gb-2006-7-11-r107
- Rep, M., Krantz, M., Thevelein, J. M., and Hohmann, S. (2000). The transcriptional response of *Saccharomyces cerevisiae* to osmotic shock. Hot1p and Msn2p/Msn4p are required for the induction of subsets of high osmolarity glycerol pathway-dependent genes. *J. Biol. Chem.* 275, 8290–8300. doi:10.1074/jbc.275.12.8290
- Ruiz-Roig, C., Viéitez, C., Posas, F., and de Nadal, E. (2010). The Rpd3L HDAC complex is essential for the heat stress response in yeast. *Mol. Microbiol.* 76, 1049–1062. doi:10.1111/j.1365-2958.2010.07167.x
- Rutherford, J. C., Jaron, S., Ray, E., Brown, P. O., and Winge, D. R. (2001). A second iron-regulatory system in yeast independent of Hct1p. *Proc. Natl. Acad. Sci. U.S.A.* 98, 14322–14327. doi:10.1073/pnas.261381198
- Sahara, T., Goda, T., and Ohgiya, S. (2002). Comprehensive expression analysis of time-dependent genetic responses in yeast cells to low temperature. *J. Biol. Chem.* 277, 50015–50021. doi:10.1074/jbc.M209258200
- Saito, H., and Posas, F. (2012). Response to hyperosmotic stress. *Genetics* 192, 289–318. doi:10.1534/genetics.112.140863
- Schade, B., Jansen, G., Whiteway, M., Entian, K. D., and Thomas, D. Y. (2004). Cold adaptation in budding yeast. *Mol. Biol. Cell* 15, 5492–5502. doi:10.1091/mbc.E04-03-0167
- Secrier, M., and Schneider, R. (2014). Visualizing time-related data in biology, a review. *Brief. Bioinform.* 15, 771–782. doi:10.1093/bib/bbt021
- Sha, W., Martins, A. M., Laubenbacher, R., Mendes, P., and Shulaev, V. (2013). The genome-wide early temporal response of *Saccharomyces cerevisiae* to oxidative stress induced by cumene hydroperoxide. *PLoS ONE* 8:e74939. doi:10.1371/journal.pone.0074939
- Shakoury-Elizeh, M., Protchenko, O., Berger, A., Cox, J., Gable, K., Dunn, T. M., et al. (2010). Metabolic response to iron deficiency in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 285, 14823–14833. doi:10.1074/jbc.M109.091710
- Shakoury-Elizeh, M., Tiedeman, J., Rashford, J., Ferea, T., Demeter, J., Garcia, E., et al. (2004). Transcriptional remodeling in response to iron deprivation in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 15, 1233–1243. doi:10.1091/mbc.E03-09-0642
- Slavov, N., and Botstein, D. (2011). Coupling among growth rate response, metabolic cycle, and cell division cycle in yeast. *Mol. Biol. Cell* 22, 1997–2009. doi:10.1091/mbc.E11-02-0132
- Slavov, N., and Botstein, D. (2013). Decoupling nutrient signaling from growth rate causes aerobic glycolysis and deregulation of cell size and gene expression. *Mol. Biol. Cell* 24, 157–168. doi:10.1091/mbc.E12-09-0670
- Tai, S. L., Daran-Lapujade, P., Walsh, M. C., Pronk, J. T., and Daran, J. M. (2007). Acclimation of *Saccharomyces cerevisiae* to low temperature: a chemostat-based transcriptome analysis. *Mol. Biol. Cell* 18, 5100–5112. doi:10.1091/mbc.E07-02-0131
- Teixeira, M. C., Mira, N. P., and Sá-Correia, I. (2011). A genome-wide perspective on the response and tolerance to food-relevant stresses in *Saccharomyces cerevisiae*. *Curr. Opin. Biotechnol.* 22, 150–156. doi:10.1016/j.copbio.2010.10.011
- Thorpe, G. W., Fong, C. S., Alic, N., Higgins, V. J., and Dawes, I. W. (2004). Cells have distinct mechanisms to maintain protection against different reactive oxygen species: oxidative-stress-response genes. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6564–6569. doi:10.1073/pnas.0305888101
- Treu, L., Campanaro, S., Nadai, C., Toniolo, C., Nardi, T., Giacomini, A., et al. (2014). Oxidative stress response and nitrogen utilization are strongly variable in *Saccharomyces cerevisiae* wine strains with different fermentation performances. *Appl. Microbiol. Biotechnol.* 98, 4119–4135. doi:10.1007/s00253-014-5679-6
- Tripodi, F., Nicastro, R., Reghellin, V., and Coccetti, P. (2015). Post-translational modifications on yeast carbon metabolism: regulatory mechanisms beyond transcriptional control. *Biochim. Biophys. Acta* 1850, 620–627. doi:10.1016/j.bbagen.2014.12.010
- van Bakel, H., Strengman, E., Wijmenga, C., and Holstege, F. C. (2005). Gene expression profiling and phenotype analyses of *S. cerevisiae* in response to changing copper reveals six genes with new roles in copper and iron metabolism. *Physiol. Genomics* 22, 356–367. doi:10.1152/physiolgenomics.00055.2005
- van den Brink, J., Daran-Lapujade, P., Pronk, J. T., and de Winder, J. H. (2008). New insights into the *Saccharomyces cerevisiae* fermentation switch: dynamic transcriptional response to anaerobicity and glucose-excess. *BMC Genomics* 9:100. doi:10.1186/1471-2164-9-100
- van Dijk, E. L., Chen, C. L., d'Aubenton-Carafa, Y., Gourvennec, S., Kwapisz, M., Roche, V., et al. (2011). XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature* 475, 114–117. doi:10.1038/nature10118
- Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63. doi:10.1038/nrg2484
- Warringer, J., Hult, M., Regot, S., Posas, F., and Sunnerhagen, P. (2010). The HOG pathway dictates the short-term translational response after hyperosmotic shock. *Mol. Biol. Cell* 21, 3080–3092. doi:10.1091/mbc.E10-01-0006
- Wu, C. Y., Bird, A. J., Chung, L. M., Newton, M. A., Winge, D. R., and Eide, D. J. (2008). Differential control of Zap1-regulated genes in response to zinc deficiency in *Saccharomyces cerevisiae*. *BMC Genomics* 9:370. doi:10.1186/1471-2164-9-370
- Wu, J., Zhang, N., Hayes, A., Panoutsopoulou, K., and Oliver, S. G. (2004). Global analysis of nutrient control of gene expression in *Saccharomyces cerevisiae*

- during growth and starvation. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3148–3153. doi:10.1073/pnas.0308321100
- Wysocki, R., and Tamás, M. J. (2010). How *Saccharomyces cerevisiae* copes with toxic metals and metalloids. *FEMS Microbiol. Rev.* 34, 925–951. doi:10.1111/j.1574-6976.2010.00217.x
- Yale, J., and Bohnert, H. J. (2001). Transcript expression in *Saccharomyces cerevisiae* at high salinity. *J. Biol. Chem.* 276, 15996–16007. doi:10.1074/jbc.M008209200
- Yamaguchi-Iwai, Y., Dancis, A., and Klausner, R. D. (1995). AFT1: a mediator of iron regulated transcriptional control in *Saccharomyces cerevisiae*. *EMBO J.* 14, 1231–1239.
- Yosef, N., and Regev, A. (2011). Impulse control: temporal dynamics in gene transcription. *Cell* 144, 886–896. doi:10.1016/j.cell.2011.02.015
- Zakrzewska, A., van Eikenhorst, G., Burggraaff, J. E., Vis, D. J., Hoefsloot, H., Delneri, D., et al. (2011). Genome-wide analysis of yeast stress survival and tolerance acquisition to analyze the central trade-off between growth rate and cellular robustness. *Mol. Biol. Cell* 22, 4435–4446. doi:10.1091/mbc.E10-08-0721
- Zaman, S., Lippman, S. I., Schnepfer, L., Slonim, N., and Broach, J. R. (2009). Glucose regulates transcription in yeast through a network of signaling pathways. *Mol. Syst. Biol.* 5, 245. doi:10.1038/msb.2009.2
- Zaman, S., Lippman, S. I., Zhao, X., and Broach, J. R. (2008). How *Saccharomyces* responds to nutrients. *Annu. Rev. Genet.* 42, 27–81. doi:10.1146/annurev.genet.41.110306.130206
- Zhao, H., Chen, J., Liu, J., and Han, B. (2015). Transcriptome analysis reveals the oxidative stress response in *Saccharomyces cerevisiae*. *RSC Adv.* 5, 22923–22934. doi:10.1039/C4RA14600J
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Taymaz-Nikerel, Cankorur-Cetinkaya and Kirdar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Advantages of publishing in Frontiers



## OPEN ACCESS

Articles are free to read,  
for greatest visibility



## COLLABORATIVE PEER-REVIEW

Designed to be rigorous  
– yet also collaborative,  
fair and constructive



## FAST PUBLICATION

Average 85 days from  
submission to publication  
(across all journals)



## COPYRIGHT TO AUTHORS

No limit to article  
distribution and re-use



## TRANSPARENT

Editors and reviewers  
acknowledged by name  
on published articles



## SUPPORT

By our Swiss-based  
editorial team



## IMPACT METRICS

Advanced metrics  
track your article's impact



## GLOBAL SPREAD

5'100'000+ monthly  
article views  
and downloads



## LOOP RESEARCH NETWORK

Our network  
increases readership  
for your article

## Frontiers

EPFL Innovation Park, Building I • 1015 Lausanne • Switzerland  
Tel +41 21 510 17 00 • Fax +41 21 510 17 01 • [info@frontiersin.org](mailto:info@frontiersin.org)  
[www.frontiersin.org](http://www.frontiersin.org)

## Find us on

